



सत्यमेव जयते

**INDIAN AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI**

I.A.R.I.6.

GIP NLK—H-3 I.A.R.I.—10-5-55—41,000

PROCEEDINGS
OF THE
ROYAL SOCIETY OF LONDON

SERIES A. MATHEMATICAL AND PHYSICAL SCIENCES

VOL 199

LONDON

Published for the Royal Society by the
Cambridge University Press
Bentley House, N.W.1

7 December 1949

Printed in Great Britain at the University Press, Cambridge
(Brooke Crutchley, University Printer)
and published by the Cambridge University Press
Cambridge, and Bentley House, London
Agents for Canada, India, and Pakistan: Macmillan

CONTENTS

SERIES A VOLUME 199

No. A 1056. 7 October 1949

	PAGE
The thermal degradation of polyvinyl compounds. I. A new type of molecular still. By N. Grassie and H. W. Melville, F.R.S.	1
The thermal degradation of polyvinyl compounds. II. The degradation of benzoyl peroxide catalyzed polymethyl methacrylates. By N. Grassie and H. W. Melville, F.R.S.	14
The thermal degradation of polyvinyl compounds. III. The effect of inhibitors and end groups on the degradation of polymethyl methacrylate. By N. Grassie and H. W. Melville, F.R.S.	24
The thermal degradation of polyvinyl compounds. IV. The thermal degradation of the methyl methacrylate copolymers with glycol dimethacrylate and acrylonitrile. By N. Grassie and H. W. Melville, F.R.S.	39
The angular distribution of γ -radiation from light nuclei. I. Experimental. By S. Devons and M. G. N. Hine	56
The angular distribution of γ -radiation from light nuclei. II. Theoretical. By S. Devons and M. G. N. Hine	73
Spectroscopic studies of low-pressure flames. II. Effective translational and rotational temperatures from CH bands. By A. G. Gaydon and H. G. Wolfhard. (Plate 1)	89
Distribution of solute atoms round a slow dislocation. By A. H. Cottrell and M. A. Jaswon	104
The absorption and dispersion of ultrasonic waves in acetic acid. By J. Lamb and J. M. M. Pinkerton	114
Optical diffraction patterns produced by bubble rafts. By J. Dyson. (Plates 2 to 7)	130

No. A 1057. 25 October 1949

Bakerian Lecture. A region of biosynthesis. By H. Raistrick, F.R.S. (Plates 8 to 11)	141
Vibration spectra of hydrocarbon molecules. II. Skeletal frequencies in certain branched paraffins. By Delia M. Simpson and G. B. B. M. Sutherland, F.R.S.	169
The use of polarized infra-red radiation in the study of doubly oriented long-chain polymers. By E. J. Ambrose, A. Elliott and R. B. Temple	183
Order-disorder statistics. II. A two-dimensional model. By C. Domb	199
Nuclear magnetic resonance at low temperatures. By J. Hatton and B. V. Rollin	222
The nature of turbulent motion at large wave-numbers. By G. K. Batchelor and A. A. Townsend	238
The analytic behaviour of Heisenberg's S matrix. By R. J. Eden	256

No. A 1058. 7 November 1949

PAGE

The organization and work of the Division of Biochemistry and General Nutrition of C.S.I.R. By H. R. Marston. (Plates 12 to 18)	273
On overpotential and the photovoltaic process at polarized electrodes. By P. J. Hillson and E. K. Rideal, F.R.S.	295
Observations on the turbulent fluctuations of a tidal current. By K. F. Bowden and J. Proudman, F.R.S.	311
The ignition and flame reactions of ethylene oxide. By F. A. Burden and J. H. Burgoyne	328
The production of mesons by electrons. By I. N. Sneddon and B. F. Touschek	352
Statistical mechanics and the partition of numbers. I. The transition of liquid helium. By H. N. V. Temperley	361
An elastic theory of dislocations. By Elizabeth H. Mann	376
The pyrolysis of methane, ethane and <i>n</i> -butane on a platinum filament. By A. J. B. Robertson	394

No. A 1059. 7 December 1949

Electromagnetic induction in a rotating sphere. By E. C. Bullard, F.R.S.	413
The flow under gravity of a swirling liquid through an orifice-plate. By A. M. Binnie and J. F. Davidson. (Plate 19)	443
The transformation of the nitrogen isotope of mass 15 into carbon and helium by bombardment with protons. By W. Cochrane and A. G. Hester	458
Unsteady compressible flow in two dimensions. By T. V. Davies	468
The electron pairing theory of the structure of conjugated hydrocarbons. By W. Moffitt	487
The kinetics of the polymerization of sarcosine carbonic anhydride. By S. G. Waley and J. Watson	499
The use of quaternions in wave-tensor calculus. By C. W. Kilmister	517
Steady flow in the laminar boundary layer of a gas. By C. R. Illingworth	533
Index	559

The thermal degradation of polyvinyl compounds

I. A new type of molecular still

BY N. GRASSIE AND H. W. MELVILLE, F.R.S.

Chemistry Department, University of Aberdeen

(Received 3 March 1949)

A survey of depolymerization reactions has been made and the importance of the degradation emphasized in addition polymerization, especially at high temperatures. In order to follow the course of the thermal degradation it is necessary to remove rapidly and measure the amount of monomer produced. In addition it is essential to measure the average size and the number of macromolecules during the course of the reaction. A molecular still for carrying out the degradation is described in which it is possible to follow the evolution of monomer continuously. The conditions necessary for the proper utilization of this type of still are described.

INTRODUCTION

A great deal of effort has gone into the precise elucidation of the mechanism of polymerization but very little is known regarding the degradation of high polymers. An understanding of both processes is necessary because in many cases a high polymer-monomer system may exist in a state of equilibrium and it is important to know something about the position of this equilibrium. In fact the size and structure of the polymer molecules will be directly controlled by the position of the equilibrium state. While in some cases it is profitable to study the equilibrium state it is more convenient to study separately the reactions of polymerization and depolymerization that lead to its establishment. This makes the study of the degradation process all the more necessary.

The first attempts to attack the problem were made by Hopff, Mark & Meyer (1929) in their experimental investigations of the hydrolysis of starch and cellulose. This work stimulated Kuhn (1930), Kuhn & Freudenberg (1930) to look into the theoretical treatment of the degradation process. For example, Kuhn showed that if every hydrolyzable bond in the molecule was equally easily severed, agreement between theory and experiment agreed closely in the initial stages of degradation. In order to explain the course of the later stages of the reaction Kuhn *et al.* suggested that hydrolysis of the bonds in the biose and triose residues took place more rapidly than in the polysaccharide; in addition it seemed necessary to assume that all terminal bonds break more quickly than those along the chain.

Even less work has been done with vinyl polymers although it has been known for some considerable time that a number of these polymers could not only be degraded but that considerable amounts of monomer could be collected in the early stages of the reaction. Similarly degradation can be brought about by mechanical shaking of the polymer, by violent agitation of solutions by the action of supersonic radiation, by the action of oxygen, peroxides, etc. These observations are all essentially qualitative in character and are not suitable for analysis so that an insight into the mechanism might be obtained.

Simha (1941) discussed in general terms in a theoretical paper how it might be possible from quantitative studies and from theoretical treatment to get some idea as to whether the degradation occurred by a mechanism generally similar to that for polysaccharides or whether, for example, preferred breaking at the ends might occur. Later Votinov, Kobeko & Marei (1942) attempted to apply Kuhn's theories to the degradation of polymethyl methacrylate and polystyrene and claimed to get good agreement in spite of the anomaly that large amounts of monomer are produced right at the start of degradation. Further, Votinov *et al.* showed that in the case of polystyrene there was a marked difference in the nature of the products depending upon whether the reaction was carried out in an atmosphere of nitrogen or of oxygen. In the latter case the products could not be repolymerized showing probably that styrene was not a product of degradation and that therefore oxidative breakdown was a contributory reaction in this particular environment. Further papers, predominantly theoretical, were published by Jellinek (1944), Tuckett (1945), Blatz & Tobolsky (1945) and Bachman (1945), but none of these papers could advance the subject for lack of satisfactory experimental data.

It is the object of this and subsequent papers to attempt to get a much closer insight into a degradation of a vinyl polymer (*a*) by choosing a system which shall be as simple as possible chemically, and (*b*) by development of more sensitive and exact experimental methods than have been used hitherto.

Of the three polymers, polymethyl methacrylate, polystyrene, and poly- α -methylstyrene, the first would appear to be most suitable for study since side reactions such as production of dimers, etc. is a minimum. Another reason for choosing methacrylate is that information is available regarding the heat of polymerization and the velocity coefficients for the radical polymerization. The discussion of equilibrium conditions is thus possible if the degradation process can be analyzed in comparable detail. Although the structure of the polymer is not known it is believed that the polymer is of the head-to-tail type. The nature of the terminal groups can be varied by altering the nature of the radical-producing reaction which initiates polymerization. The experimental conditions to cut down complications are clearly that the polymer should be heated *in vacuo* in the form of a thin film so that monomer can escape readily. The problem of following the rate of evolution of monomer will be dealt with in detail later. The average size of residual polymer molecules must be determined by a number average method since only then can a reliable measure be obtained of the number of bonds broken in the degradation reaction.

Votinov *et al.* collected the monomer vapours and weighed them from time to time so that their values for the rate of reaction are averages over long periods of time. An improvement on this procedure would be to weigh the degrading polymer *in vacuo* while the reaction was proceeding but there are considerable difficulties in weighing accurately the tray in a molecular still while maintaining the temperature at a constant value.

The normal type of molecular still consists of a heated surface, from which the substance to be distilled is evaporated, and a cooled surface only a few millimetres from the heated surface, on which the purified substance is condensed. The whole is contained in an envelope under high vacuum. Whenever a molecule evaporates

from the hot surface it will move directly to the cold surface and condense so that the pressure of the material in the still would be extremely small. The only possible methods of finding the rate of the reaction would be either to measure the increase in weight of the cooled surface or the decrease in weight of the heated surface, both difficult matters.

If, however, the cooled surface were moved away from the heated surface so that evaporated molecules did not pass directly to it but first suffered several collisions with the walls of the containing vessel then a small pressure would be set up in the still and a determination of this pressure would give a measure of the rate of the reaction. This is the principle on which the apparatus was built, the heated and cooled surfaces being separated by a mercury vapour diffusion pump, the pressure inside the still being measured by a Pirani thermal conductivity gauge.

THE MOLECULAR STILL

Glasswork

A diagram of the glasswork of the apparatus is shown in figure 1, where 1 is the molecular still which is constructed from a vacuum desiccator.

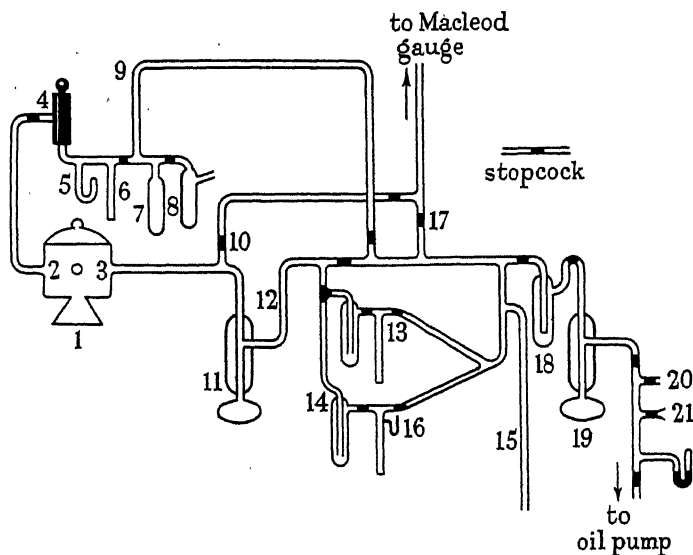


FIGURE 1. Glasswork of molecular still.

Half way up the body of the desiccator, and symmetrically round it, four holes were bored and into them four B 24 standard sockets were sealed with vacuum wax. Two of these were used to take in electrical leads to supply the electrical components in the still. 2 leads to a part of the apparatus only used when calibrating the still, 4 is a needle valve, 5 a manometer, 6 a calibrated capillary, and 7 and 8 two containers used for distilling monomer under vacuum before using it to calibrate the apparatus. This part of the apparatus is evacuated through 9 thus bypassing the still.

The still itself is evacuated through 3. The first branch on this line, namely at 10, leads to a Macleod gauge. 11 is a mercury vapour diffusion pump and behind it 12 is

the main vacuum line, but there are two identical bypasses 13 and 14 through which the products of the degradation may be led, collected and measured. Each contains a liquid air trap for collecting and a calibrated capillary for measuring the monomer produced. In addition, sealed to the top of one of the capillaries was a manometer 16 of very small volume whose use will be described later. 15 is a barometer tube and 17 leads to the Macleod gauge. The main vacuum line then passes through a liquid air trap (18), another mercury vapour diffusion pump (19) and finally to a rotary oil pump, 20 and 21 being air inlets.

Heater and heating surface

In this particular problem the whole purpose of the molecular still conditions would have been defeated if the monomer molecules had been recaptured and repolymerized inside the mass of the degrading polymer. For this reason the heating block was designed so that the requisite amount of polymer was spread as a thin film over the heated surface. The area had to be kept within reasonable limits to ensure uniformity of heating.

The heater unit is shown in figure 2. The heater itself was made from copper sheet, 3 mm. thick, and was 25 mm. square. Two such pieces of copper were used with the heating element clamped between. This heating element consisted of Nichrome wire wound on a piece of mica 20 mm. square, with a similar piece of mica above and below insulating it from the copper. The ends of this wire were soldered to thin strips of copper sheet to which in turn were soldered the wires leading out of the still to the source of current. The two copper squares and the heating element were held together by two screws placed at diagonally opposite corners.

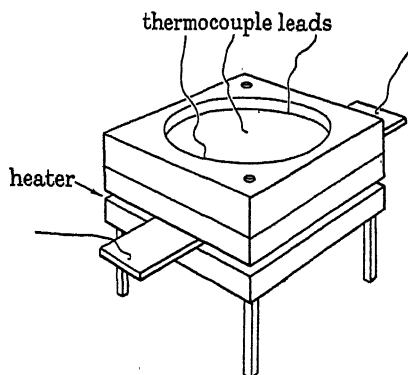


FIGURE 2. The heater unit.

The actual heating surface was a copper tray of exactly the same area as the heating block and attached to the heating element. It was cut from a sheet of copper 6 mm. thick. In the middle of this sheet of copper and 23 mm. in diameter a circular hole was bored to a depth of 5 mm. to contain the polymer.

Thermocouples

The layer of polymer in the tray must be thin, so it is difficult to measure the temperature of the polymer simply by dipping the thermocouples into it. The thermocouple leads (one common copper and two constantan, one at the edge and one at

the centre of the tray) were therefore hard soldered on to the bottom of the inside of the tray.

This whole heater unit was supported on a stretcher made from silica rod, which was itself supported on the shelf of the desiccator.

Pirani gauge

The Pirani gauge, figure 3, was built into the apparatus underneath the heater unit. To measure the actual pressure in the still and to register quickly a changing pressure this gauge had to be placed in the vessel containing the vapours and not in a side arm as is usually done when measuring static pressures.

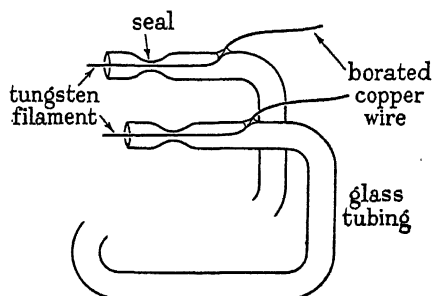


FIGURE 3. The Pirani gauge.

The gauge consisted of a 3 cm. length of tungsten single coiled electric lamp filament, the ends of the filament being mounted on a rigid glass framework with suitable leads to an external resistance bridge.

Electrical circuits

Each of the three main components inside the still, namely the heater, the thermocouples and the Pirani gauge, had an electrical circuit associated with it.

The heater circuit. The heater current was taken from a 24 V transformer. This current was passed through two variable resistances in parallel, and through the heater which was in series with them. The thermo regulator was also in series with one of these resistances so that it could be switched off and on as the temperature of the heating tray fell below and rose above the required temperature.

The temperature regulator was a Cambridge potentiometer model.

The thermocouple circuit. The copper heating block in the still constitutes the hot junction while the cold junction was a thermos flask of melting ice. The constantan leads each passed through a switch before going to the millivoltmeter so that the temperature could be obtained from either thermocouple by depressing the appropriate switch. In this way it was possible to verify that the temperature was constant over the surface of the tray.

From the copper and one of the constantan leads, two other leads were taken to the temperature regulator, one of them through a switch. Provided one thermocouple was used to obtain the temperature and the other to work the thermoregulator then the regulator had no effect upon the temperature reading.

The Pirani gauge circuit. This was the standard Pirani gauge circuit consisting of a simple Wheatstone bridge.

The source of current was an ordinary 2 V accumulator discharging through a switch and a variable resistance. A voltmeter measured the voltage across the bridge while in use.

The bridge was constructed with two 50Ω resistances as two of the arms, the Pirani as the third and a resistance, which was made withappings such that any value from 0 to 300Ω (in steps of 10) could be obtained, as the fourth.

CALIBRATION OF THE APPARATUS

Rate measurement

With the Pirani gauge circuit described in the last section it is known that the expression $(V^2 - V_0^2)/V_0^2$ (V is the voltage across the bridge when balanced at given pressure and V_0 when balanced at zero pressure), which is the fractional increase in energy, is linear with pressure up to a pressure of about 0.1 mm. of mercury. Accordingly this expression should be proportional to the rate of production of monomer in the still. To verify this point, monomer was allowed to flow at various rates through the needle valve (4, figure 1) and through the still, being then collected in the liquid air trap (13, figure 1), distilled into the calibrated capillary, and measured. Values of V and V_0 were noted and the expression $(V^2 - V_0^2)/V_0^2$ plotted against rate of flow of monomer through the apparatus in ml. per hour. Results are shown in figure 4.

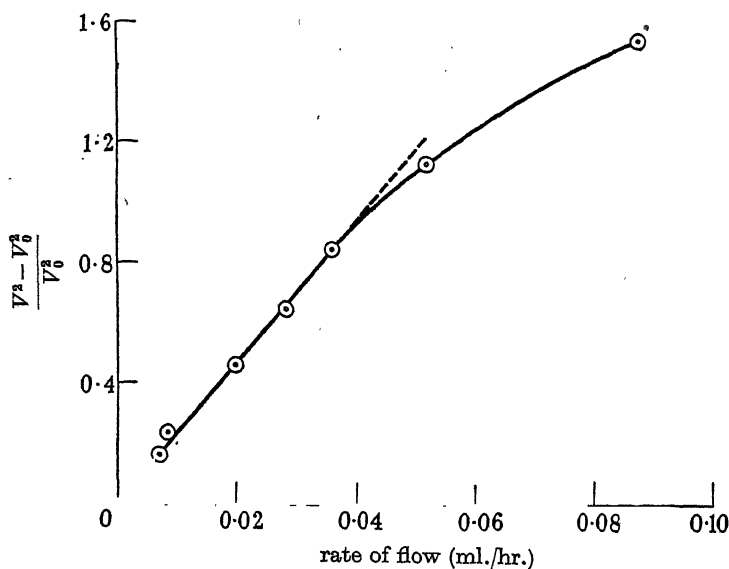


FIGURE 4. Pirani gauge calibration curve.

From figure 4 it can be seen that the relationship is linear up to a rate of flow of 0.04 ml./hr. and $(V^2 - V_0^2)/V_0^2 = 0.95$. Assuming that a measurement may be made in 10 sec., this corresponds to 10^{-4} g. 10^{-5} g. could be measured corresponding to one tenth of the full rate of monomer production. Higher rates could of course be measured by suitable calibration of the gauge.

Owing to the close proximity of the heater and the Pirani gauge the thermal radiation from the heater affected the gauge to a marked extent, the voltmeter reading

being depressed by as much as 0.05 V on raising the temperature to 320° C. A shield, made from thin copper sheet, was suspended between the two, no appreciable effect being then noticed up to 350° C. Figure 5 shows values for the voltage across the bridge when balanced, on raising the temperature of the copper block, with and without the copper shield in position. From the graph it can be seen that the shield eliminates the trouble.

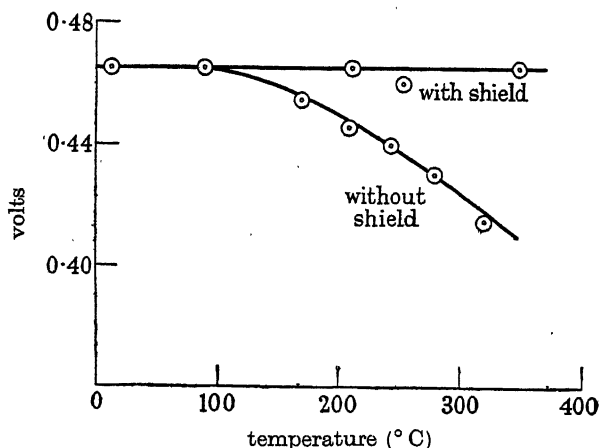


FIGURE 5. Effect of copper shield on Pirani gauge readings.

Constant pumping speed

The pressure produced inside the still will depend on the rate of pumping of the mercury vapour diffusion pump (11, figure 1). It was therefore essential to find what variables affected pumping speed especially with monomer production at its maximum value on the linear portion of the Pirani gauge scale.

The variables are rate of heating and rate of flow of cooling water. First of all, using a rate of flow of cooling water of about 2 l./min., the pressure of coal gas passing to the pump burner was varied and between gas pressures of 23 and 40 mm. of water the pressure in the still remained strictly constant. Below and above these values the pressure in the still increased. Change in water pressure made no difference to the lower limit but if the water pressure was increased to 4 l./min. then the upper gas pressure limit was raised to 56 mm. It appears that the lower limit of 23 mm. gas pressure gives the lowest possible pumping rate which can cope with this rate of flow of monomer and carry away all monomer which reaches the pump by diffusion. The higher gas pressure limit, which is different for each cooling water rate, is probably due to the fact that the mercury is not properly condensed.

As a result of these preliminary experiments, the gas pressure was set between 30 and 40 mm. before each experiment and the water rate above 2 l./min.

Thermocouples

The thermoregulator used for controlling the temperature of the heating block was a Cambridge potentiometer model and was used with copper-constantan thermocouples.

Owing to the necessity of taking leads through vacuum-tight seals, the thermocouple circuits necessarily included sections of borated copper wire, which might have upset the thermocouple readings. For this reason the circuit was exactly duplicated outside the apparatus and the e.m.f. measured with the copper tray in baths at various temperatures, measured by standard thermometers. The e.m.f.'s lay exactly on the standard curves, at least up to 350° C.

Heater circuit and thermoregulator

Owing to the small heat capacity of the copper heating block, accurate temperature control was very difficult and whereas the thermoregulator could keep the temperature of a relatively large furnace to within a fraction of a degree of the required temperature, it could, at its best, only keep the copper block within about $\pm 2^\circ$.

This $\pm 2^\circ$ drift of temperature raised the problem of getting accurate values of reaction rate at any required temperature. The conditions of control are such that the temperature rises very quickly from -2° to $+2^\circ$ but falls back very slowly. Readings were taken as the temperature was falling and its value noted at the required temperature.

EXPERIMENTAL

A number of variables that may vitiate results still remain. The methods of control are discussed below.

Form of the polymer

It seemed that for uniform heating of polymer samples the polymer would have to form a compact layer on the bottom of the tray and the best method of achieving this seemed to be to use the polymer in the form of a very fine powder. With the polymer in this state it was found that rate measurements could be reproduced accurately. With larger particles the rate of the reaction might depend upon grain size—the monomer molecules produced finding it more difficult to escape from the larger grains. In this way the degradation rate might be some function of the surface area of the grains.

Experiments were carried out under identical conditions using samples of a benzoyl peroxide catalyzed polymethyl methacrylate having a molecular weight of 179,000. 20 to 80 mesh samples were used and the temperature of the tray was 260° C. Only 0.0281 g. polymer was used so that the thickness of the sample on the tray was very small, each particle being in direct contact with the tray. $(V^2 - V_0^2)/V_0^2$ was plotted in each case against size and the results are shown in figure 6. It can be seen from the figure that the rate is independent of particle size for particles smaller than about 50 mesh but that it does fall away for larger particles. This diminution in rate with the larger particles may be due to the monomer molecules being trapped inside the polymer or it may be due to the fact that the larger particles are not uniformly heated to the temperature of the tray. Polymer is a bad conductor of heat and the side of the particles away from the tray may be at a temperature considerably less than 260° C. As already explained all these experiments were done with a monolayer of polymer particles so if this is the explanation of the fall away in the rate of

degradation of large particles then with more than a monolayer of small particles in the tray the rate will be abnormally small. The small amounts of polymer contained in monolayers, namely about 0.025 g., were not enough for molecular weight measurements of the residual polymer and therefore it was necessary to make certain, if say 0.3 g. were present in the tray, that the whole of that 0.3 g. was at the required temperature and therefore degrading at the same rate. It was also found, in the case of low molecular weight samples, that if the size of the polymer particles was much larger than 80 mesh then each particle formed a bubble on heating and the volume of the polymer became very great.

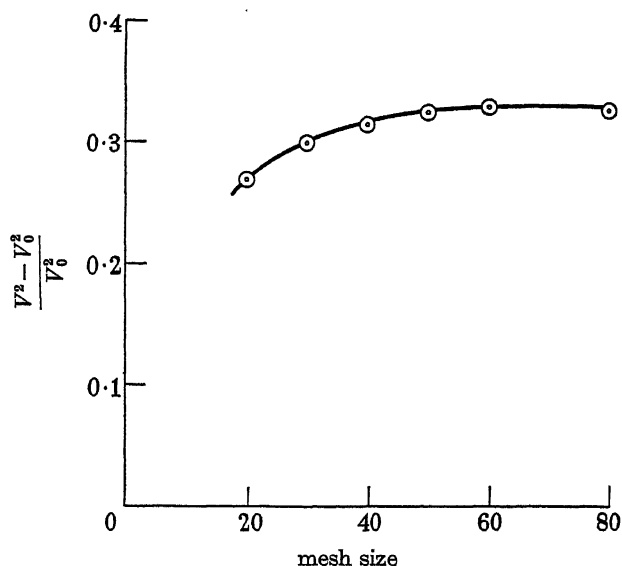


FIGURE 6. Effect of polymer grain size on rate of degradation.

Effect of layer thickness

In order to find if all the polymer was degrading, a series of experiments was carried out at the same temperature with different amounts of the same polymer. The value $(V^2 - V_0^2)/V_0^2$ was plotted against the weight of polymer in the tray. If the curve or part of it were a straight line then it would be evident that under the conditions of the straight part all the polymer was at the temperature of the tray.

Such experiments were carried out at 260° C and the results are shown in figure 7 (full line).

It will be seen that the polymer was heated satisfactorily provided there was not more than a maximum of 0.034 g. in the tray. If there was more than this amount then the upper layers had a temperature lower than that of the tray.

One solution of the problem was the use of bigger trays, but it is difficult to maintain constancy of temperature over a large surface.

At this stage the catalytic effect of copper was being investigated to make sure that the tray itself had no such effect upon the course of the reaction. Copper powder was mixed intimately with the polymer and it was found that there was no rate

change, provided the weight of polymer used was less than 0.04 g., and hence there was no catalytic effect. It was discovered, however, that if the normal rate was on the *BC* part of the curve then there tended to be an unreproducible increase in rate when copper was present. Since there was no catalytic effect this effect must be due to the copper present producing more uniform heating of the polymer.

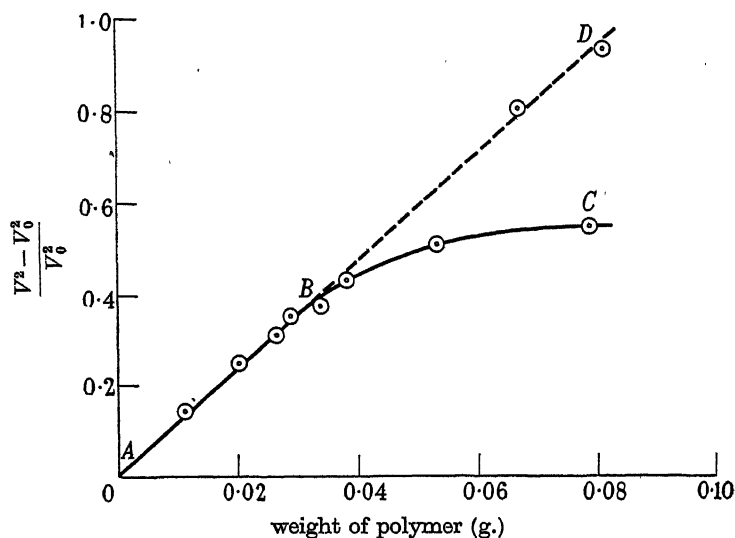


FIGURE 7. Effect of layer thickness on rate of degradation at 260°C.

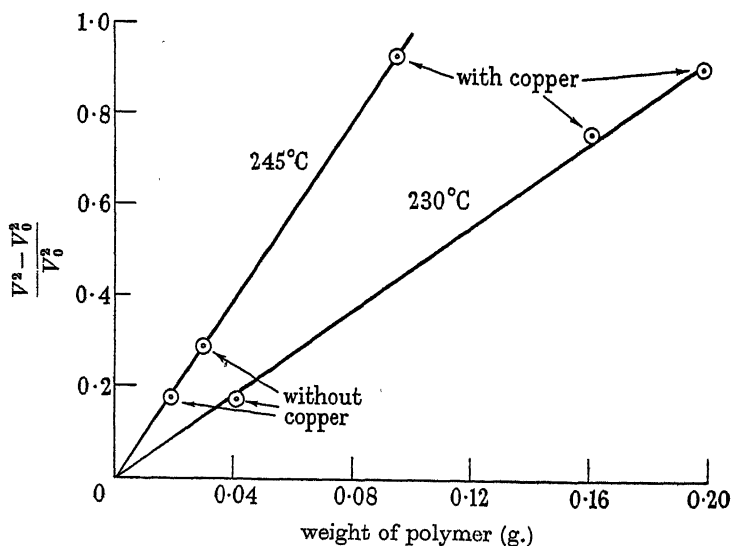


FIGURE 8. Rates of degradation at 230 and 245°C, with and without copper.

After a number of trials it was found that the most satisfactory solution of the problem consisted in covering the polymer completely with copper powder, such that the copper formed a continuous layer over it, and not by simply mixing the two, which only placed some of the polymer farther away than ever from the heating

surface. Points obtained in this way are shown on the *BD* portion of figure 7. Results obtained at other temperatures are shown in figure 8. Experiments also at 210° C showed that the amount of polymer in the tray could be increased to at least 0.5 g. and the linear relationship maintained. 0.5 g. of polymer was the maximum amount which the tray could hold.

The copper powder used in these experiments was always obtained by the reduction of B.D.H. cupric oxide in a stream of coal gas. After reduction it was ground quickly in a mortar and passed through an 80-mesh sieve. That which passed through the sieve was used and the rest discarded. Until used it was stored in a tightly stoppered bottle owing to its relatively rapid oxidation in air.

Rate measurement

On placing a sample of polymer in the heating tray and slowly raising the temperature, a plot of Pirani gauge volts as a function of time is obtained as shown in figure 9. The surprising feature is the existence of two maxima, and the problem then is to decide what points give a measure of monomer production.

It seemed hardly likely that the material evolved, which was causing these two maxima, was a product of degradation (i.e. monomer) and to give some indication of its nature its vapour pressure was measured. Samples of the vapours evolved up to *A* in figure 9 were collected and distilled into the calibrated capillary and in every case the vapour pressure of the product was about 40 to 50 mm. whereas that of pure methyl methacrylate at room temperature is only about 20 mm. Samples taken after *A* showed only the vapour pressure of monomer.

The explanation which was put forward to explain the shape of these degradation curves is as follows. The two maxima are due to the presence of solvents in the polymer, solvents like methyl alcohol and acetone having been used in the preparation of the polymer. Under vacuum and on gentle heating the solvents adsorbed on the surface of the polymer are evolved and cause the first maximum. This is further borne out by the fact that this maximum may be decreased by keeping the polymer for a prolonged period of time *in vacuo* before heating. The second maximum only occurs at a very much higher temperature and the solvent molecules causing it were originally trapped inside the polymer and are evolved as it fuses to a viscous mass. It was found that for polymer samples of different molecular weights and which have therefore different softening points, the temperature at which these volatiles were evolved and the maximum occurred was closely related to the softening point.

To obtain a value for the initial rate of degradation, heating was continued until the curve became linear, in this case at *A*, the time at which it first reached the required temperature having been noted by an arrow on the curve. The linear portion was then produced back (dotted line) to the time at which the required temperature was first reached and the value of the rate at this point was taken as the initial rate. This procedure has been adopted in all the experiments described in this and the following papers. The most conclusive way, however, of showing the validity of this method will become evident in a later paper. Briefly it is as follows—the rate after all volatiles have been eliminated, that is after point *A* in figure 9, is given simply

by the values of V and V_0 . If these rate values, $(V^2 - V_0^2)/V_0^2$, are plotted against percentage degradation to monomer and produced back to zero degradation it is found that the initial rate is always, within experimental error, the same as the rate obtained by the method described above.

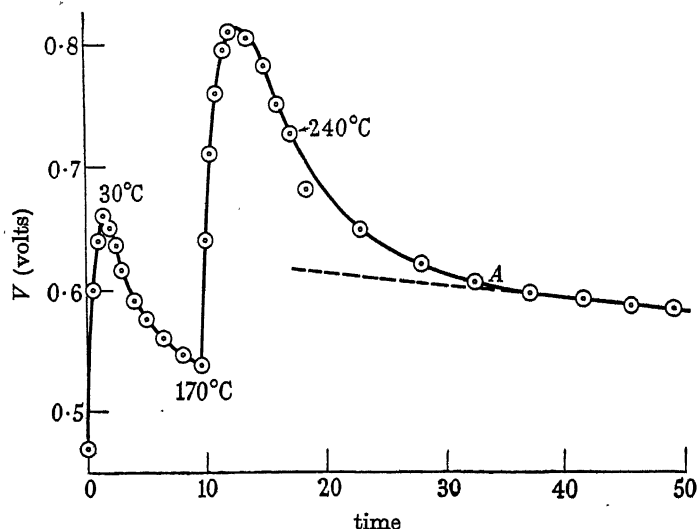


FIGURE 9. Elimination of volatiles and determination of initial rate of degradation.
 $V_0 = 0.47$. 0.1529 g. polymer.

Production of the polymer in the form of a powder

The importance of the form of the polymer to be degraded has already been discussed. It was shown experimentally that provided the polymer was present as a fine powder and provided it was covered by a layer of copper powder then results could be made reproducible because it was certain that under these conditions all of the polymer was being maintained at the correct temperature, and all the monomer which was produced in the reaction was able to escape from the polymer without side reactions taking place.

First of all it was found to be essential to start with the polymer in pure acetone solution. If the polymer had been prepared in monomer solution then the whole is simply poured into two or three times its volume of acetone. If, however, the polymer is prepared in a solvent other than acetone, it must be precipitated with methyl alcohol, roughly dried to get rid of excess solvents, and then redissolved in acetone. It is fortunate that it is only low methyl methacrylate polymers which are prepared in solution because these low polymers are much more readily soluble in acetone, and precipitation and subsequent redissolving in acetone present little difficulty. Solution of large quantities of high polymers in acetone, however, is a much more difficult matter.

The acetone solution of polymer is then slowly poured into about three times its volume of methyl alcohol with vigorous stirring. The polymer is precipitated as a white solid. As more and more of the solution is added the polymer precipitated tends to become more and more gelatinous owing to the rising concentration of acetone in

the methyl alcohol which then acts as a partial solvent or incomplete precipitant. Only about half of the solution is added; stirring is continued for 2 or 3 min. and the polymer allowed to settle. The supernatant liquid, a mixture of methyl alcohol and acetone, is then decanted and discarded. Fresh methyl alcohol is added to the polymer precipitate and almost at once it changes to an amorphous mass since complete precipitation has taken place. Vigorous stirring is continued to break up the polymer, thereby allowing the methyl alcohol to penetrate and complete the precipitation.

On allowing the polymer to settle the supernatant liquor is poured off. It consists of almost pure methyl alcohol and so it may be used for the initial precipitation of the polymer in the other half of the acetone solution.

The excess methyl alcohol is roughly squeezed out of the polymer, which is pulled out into small fibrous pieces. It is then placed in watch glasses to dry slowly in the air. It is necessary that the polymer should not be dried in an oven at this stage otherwise the small amounts of methyl alcohol present will act as a partial solvent at the elevated temperature and the pieces of polymer will fuse to very hard particles which are difficult to grind to a fine powder.

As soon as the polymer has dried sufficiently (usually in about 4 to 5 hr.) it may be ground quite easily, most of it passing through an 80-mesh sieve. Drying should not be continued for too long since the polymer becomes much more difficult to grind.

Once the polymer has been powdered it may then be thoroughly dried by heating in a steam oven for 24 hr. If it is not finally dried in this way difficulty is sometimes encountered when it is placed in the still and the pressure reduced. Under these conditions the volatiles are evolved readily and tend to carry the polymer out of the heating tray.

REFERENCES

- Bachman, G. B. 1945 *J. Organ. Chem.* **12**, 108.
Blatz, P. J. & Tobolsky, A. V. 1945 *J. Phys. Chem.* **49**, 77.
Hopff, H., Mark, H. & Meyer, K. H. 1929 *Ber. dtsh. chem. Ges.* **62**, 1103.
Jellinck, H. G. 1944 *Trans. Faraday Soc.* **40**, 266.
Kuhn, W. 1930 *Ber. dtsh. chem. Ges.* **63**, 1503.
Kuhn, W., Freudenberg, K. 1930 *Ber. dtsh. chem. Ges.* **63**, 1510.
Simha, R. 1941 *J. Appl. Phys.* **12**, 569.
Tuckett, R. F. 1945 *Trans. Faraday Soc.* **41**, 351.
Votinov, A., Kobeko, P. & Marei, F. 1942 *J. Phys. Chem. U.S.S.R.* **16**, 106.

The thermal degradation of polyvinyl compounds

II. The degradation of benzoyl peroxide catalyzed polymethyl methacrylates

BY N. GRASSIE AND H. W. MELVILLE, F.R.S.

Chemistry Department, University of Aberdeen

(Received 3 March 1949)

The thermal degradation of polymethyl methacrylates prepared by means of benzoyl peroxide has been investigated. The rate of monomer evolution is independent of molecular size for a given weight of polymer except that it falls off when the molecular weight approaches 500,000. Similarly at sufficiently low molecular weights the number of molecules in the tray of the still continuously decreases while the average size remains constant. These facts point to the mechanism's involving the fission of the molecule into radicals, the fission occurring at the end of the molecule. Thereafter the large radical rapidly undergoes degradation, monomer units being evolved until the radical is completely removed. With the larger radicals there is not time for complete degradation; instead these interact by disproportionation and thus become stabilized.

INTRODUCTION

As will be shown in the following papers the mode of preparation of polymethyl methacrylate has an important influence on the course of its degradation. It is therefore essential in studying some of the features of these reactions to ascertain the general structure of the polymer present. In the present paper the polymerization of the monomer has been initiated by the thermal decomposition of benzoyl peroxide in which it is believed that both phenyl and benzoate radicals start off the growth of the polymer.

PREPARATION OF POLYMER SAMPLES

Purification of monomer

Unless otherwise mentioned the monomer was purified by repeated shaking with sodium hydroxide solution until the sodium hydroxide layer was colourless. The monomer was then shaken thoroughly two or three times with distilled water to remove sodium hydroxide and it was finally allowed to stand overnight over calcium chloride.

Purification of benzoyl peroxide

The 30 to 40 % of water which is present in commercial benzoyl peroxide was removed in the following way: First the impure benzoyl peroxide was dissolved completely in a minimum volume of chloroform without heating, leaving the water as a supernatant liquor. The mixture was thoroughly shaken in a separating funnel and after allowing to settle the chloroform layer was run off. On adding methyl alcohol to this chloroform solution the benzoyl peroxide was precipitated as pure white crystals. Addition of methyl alcohol was continued until precipitation was complete and the precipitate was filtered on a Buchner funnel. Suction was continued until

the crystals were quite dry, this purified benzoyl peroxide was then stored in a desiccator over calcium chloride.

Polymer of mol.wt. = 725,000

0.1 % catalyst was used and the polymerization carried out at 60° C. 0.1 g. benzoyl peroxide and 100 ml. of methyl methacrylate were placed in a 500 ml. round-bottomed bolt-head flask which was loosely stoppered and placed in a thermostat at 60° C. After about 200 min. the liquid had become viscous and was poured into twice its volume of acetone. (The subsequent powdering of the polymer has been discussed at the end of the previous paper.)

An alternative method of polymerizing in absence of air was also employed. A sample of the purified monomer was further purified by distillation at 35° C under reduced pressure, the first and last 25 ml. of 100 ml. batch being discarded. This was then thoroughly degassed and polymerized *in vacuo* under the same conditions of temperature and with the same concentration of catalyst as described above. The polymer so produced had a number average mol.wt. of 650,000, thus being practically the same as that prepared in presence of oxygen. Later it will be shown that the rate of degradation of this polymer and of the 726,000 polymer made from monomer purified as at the beginning of this section, were exactly the same, so it may be assumed that the method of purification of monomer is sufficient, at least for preparing polymers of this type.

Polymer of mol.wt. = 179,000

This polymer was prepared using 0.5 % catalyst at 80° C, the experimental details being the same as for the 725,000 polymer above.

Polymer of mol.wt. = 94,000

With the high catalyst concentrations and temperatures required to produce a polymer of this molecular weight the polymerization reaction in pure monomer proceeds very quickly and is difficult to control. Part of this difficulty is probably due to the fact that as the liquid becomes viscous the large amounts of heat evolved cannot be carried away quickly enough by convection currents. The result is that local heating takes place which in turn makes the reaction proceed at an even greater rate and results in less uniform polymers.

Polymerization was therefore carried out in solution as follows: 100 ml. of methyl methacrylate, 500 ml. ethyl alcohol, and 1 g. benzoyl peroxide were placed in a 2l. bolt-head flask fitted with a reflux condenser. The whole was heated for 4 hr., the solution remaining clear throughout, the polymer being quite soluble in boiling ethyl alcohol. On allowing the flask and its contents to cool the polymer precipitated out as a solid white layer on the bottom of the flask. The supernatant liquor was poured off and the excess solvent squeezed out of the polymer which was then roughly dried in an oven. Finally the lump of polymer was broken up and dissolved in hot acetone.

Polymer of mol.wt. = 44,300

50 ml. methyl methacrylate, 450 ml. benzene, and 2 g. benzoyl peroxide were heated under reflux for 2 hr., cooled and poured into 2l. of methyl alcohol. The

polymer precipitated as a gummy mass which was dried in a steam oven. It was then broken up and dissolved in acetone.

Other samples

Besides these samples mentioned above a range of samples was obtained from Imperial Chemical Industries Ltd., Plastics Division, and these were used in rate measurement experiments only.

The preparations of these samples are summarized in table 1.

TABLE 1. PREPARATIONS OF POLYMER SAMPLES

conditions of preparation	percentage of catalyst	temp. (° C)	M_n
10 % in benzene	4.0	80	36,000
	2.0	100	71,000*
	2.0	80	75,000*
	0.7	100	122,000*
pure monomer	0.5	100	190,000*
	0.5	100	239,000*
	0.5	100	356,000*
	1.0	80	500,000

* Interpolated from viscosity mol.wt.

MOLECULAR WEIGHT MEASUREMENTS

The weakness of all work to date on the degradation of high polymers has been the fact that molecular weights have always been measured by the viscosity method and not osmotically. The viscosity of a high polymer solution depends not only upon the number of molecules present but also upon the size of these molecules and since there are wide differences of opinion as to exactly how the viscosity depends upon the size of the molecule it is impossible to count the number of molecules present in any sample by this method. Inability to count the exact number of molecules present means that even the simplest points about reaction mechanism cannot be clarified with certainty. If, for example, each molecule broke once, the counting of molecules by osmometry would indicate a fall in molecular weight to half the original value. Viscosity measurements, however, could do no more than simply indicate a fall in molecular weight.

In this work viscosity measurements were only made occasionally to give a rough indication of the course of the reaction, and during the experiments with 1, 4-diamino-anthraquinone to decide between possible mechanisms. In no case were molecular weights calculated from these viscosity measurements.

All exact values of molecular weight which are mentioned throughout this work were measured osmotically by the dynamic method using the improved Fuoss & Mead (1943) type of osmometer and bacterial cellulose membrane due to Masson & Melville (1946, 1949).

THEORETICAL

The choice of this particular group of polymers for the initial study of the reaction followed from the fact that it was easy to prepare samples of such polymer having a wide range of molecular weights, and since the structure of all the samples is identical

except for the numbers of methyl methacrylate units in the chains, then any difference in the degradation characteristics can only be due to the difference in chain length.

The great difficulty in trying to analyze the mechanism of degradation by means of any theory to date has been the lack of experimental evidence based essentially on values of the number average molecular weight of the polymer residue. If conditions are sufficiently ideal it should be possible, by a study of the curves of molecular weight of residue against percentage degradation, to decide which is the operative mechanism. The experimental conditions, described in the last paper, are such that all volatile substances are removed from the degrading polymer as soon as they are formed. Hence only the molecular weight of the non-volatile residue is measured.

Kuhn's (1930) random breaking theory, which was put forward more or less successfully to explain the hydrolytic degradation of high polymers, was insufficient as such to explain the thermal degradation of high polymers, mainly due to the fact that large amounts of monomer were produced even in the initial stages of the reaction. It still seemed possible that this might make up part of the reaction, combined with some other mechanism which would produce the necessary amounts of monomer.

Simha (1941) suggested that the mechanism whereby the necessary monomer was produced consisted simply of preferred or even exclusive breaking at the ends of the chains. Blatz & Tobolsky (1945) considered polymerization to be a stepwise process and they proposed that depolymerization was the exact reverse of polymerization. According to these workers then, depolymerization would also be a stepwise process and in fact this is exactly the same mechanism as Simha's exclusive breaking at the ends of the chains.

This stepwise theory of polymerization would suggest that the initial polymer which is formed in the polymerization reaction must have a low molecular weight which will increase with time. It has, however, been found that initial polymer has essentially the same molecular weight as polymer produced later in the reaction in the absence of the gel effect and the accepted explanation of this is that the actual propagation step of polymerization is extremely rapid compared with initiation and termination, so that once a polymer chain is initiated it grows very rapidly to its full size. Accepting this then as the mechanism of the polymerization reaction, a reverse polymerization would consist of the thermal fission of the polymer molecule into radicals, followed by a rapid and complete disintegration of these polymer radicals to monomer.

From these considerations, three possible extreme degradation mechanisms present themselves although any combination of the three seems possible. These three extreme mechanisms may be described as *E*, random breaking; *F*, stepwise degradation; and *G*, reverse polymerization.

Each of these mechanisms gives a distinct type of molecular weight of residue—percentage degradation to monomer curve as may be seen in figure 1.

If the mechanism is random breaking then the molecular weight would fall rapidly to a very low value before any appreciable amounts of monomer were produced and the form of the curve would follow the path *ADC*. If, on the other hand, stepwise degradation were the mechanism then the reaction would follow *AC* since the fall

in molecular weight would be proportional to the amount of monomer produced. Finally, if reverse polymerization were the mechanism, there would be no fall in molecular weight throughout the whole reaction, which would follow the path AB . According to this mechanism, if degradation in a molecule were initiated by molecular fission, then that molecule would disintegrate completely to monomer leaving no involatile residue, so that the average molecular weight of the undegraded molecules would remain unchanged.

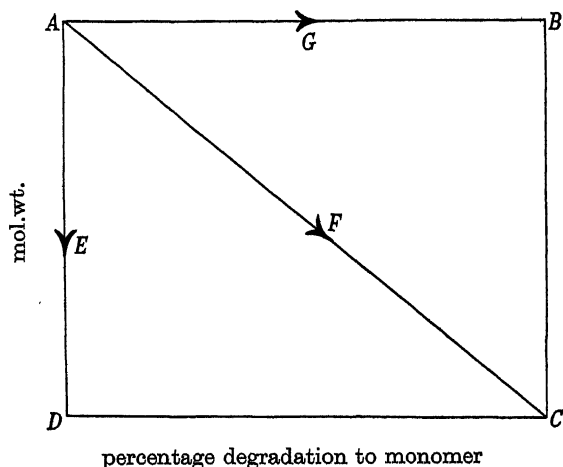


FIGURE 1. Theoretical curves for three possible extreme degradation mechanisms.

EXPERIMENTAL AND DISCUSSION

Effect of degradation on molecular weight

Curves of molecular weight of residue against percentage degradation to monomer were obtained for five samples. These had molecular weights of 44,300, 94,000, 179,000 and 725,000 and 650,000.

The percentage degradation to monomer was determined by collecting the monomer produced during the reaction in the liquid air trap (13, figure 1 of the preceding paper) and distilling it, after the reaction, into the calibrated capillary. This gave the volume of monomer, which could be converted to weight, the density of monomeric methyl methacrylate being known to be 0.945. 5 % of the initial weight of polymer was always subtracted from the weight of monomer collected to allow for the volatiles produced early in the reaction and which were not a product of degradation.

The results obtained in these experiments are summarized in figure 2.

With 44,300 polymer there was no fall in molecular weight up to at least 65 % degradation and therefore the only mechanism by which this reaction could have proceeded was 'reverse polymerization'. Each molecule, after being initially rendered unstable must disintegrate completely to monomer leaving no involatile residue. This series of experiments, however, gives no indication how the chain breaks initially; whether, for example, each bond in the chain has the same possibility of initial rupture or whether the ends of the chains are the vulnerable points.

As the molecular weight of the initial polymer is raised, the mechanism ceases to be strictly reverse polymerization in the later stages of the reaction and the molecular weight falls. From figure 2 it may be seen that the degradation of 94,000 polymer is only a reverse polymerization up to about 30 % degradation. With 179,000 polymer the molecular weight falls after about 10 to 15 % degradation while with 725,000 polymer there appears to be an immediate fall in molecular weight.

Considering the fact that the degradation of 36,000 polymer is exclusively reverse polymerization and that the 94,000 and 179,000 samples show the characteristics of this mechanism in the initial stages, it is inconceivable that some totally different reaction is proceeding in the later stages of the degradation of these samples and throughout the whole degradation of 725,000 polymer.

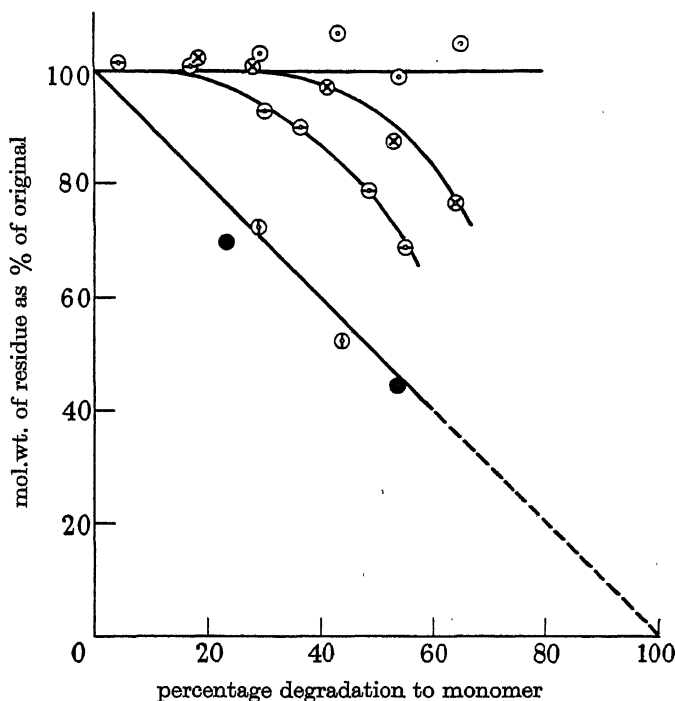


FIGURE 2. The effect of degradation on the molecular weight of a range of polymethyl methacrylate samples. ● 44,300; ⊗ 94,000; ⊖ 179,000; ⊕ 650,000; ● 725,000.

The degradation curve for 725,000 polymer would appear to point to stepwise degradation, described earlier in this paper. In this reaction one molecule of monomer would be evolved at a time from the ends of the chains and the fall in molecular weight would be proportional to the number of monomer molecules evolved. It is obvious, however, that the same number average molecular weight will result whether one monomer is evolved from each of n polymer molecules or whether n monomers are evolved from one polymer molecule. From this it would appear that the main reaction in 725,000 polymer is in fact reverse polymerization but that every degrading chain is stabilized before it disappears completely and becomes once more a stable polymer molecule, but of shorter chain length.

The degradation of the 94,000 and 179,000 polymer samples must be intermediate between these two extremes. The probability of termination seems to be just about the same as the probability of the concurrent removal of 1000 monomer units, in which case some 94,000 molecules would disintegrate completely whereas others would be prematurely terminated. The probability of complete disintegration would become smaller the higher the molecular weight and thus it is not surprising that the molecular weight of the higher molecular weight samples falls away at an earlier stage in the reaction.

From these considerations it might be expected that the molecular weight of the 94,000 and 179,000 samples would decrease slightly even in the initial stages, but the detection of such a decrease is difficult. Those degradation characteristics are complicated to some extent by the fact that the polymer samples always have a wide molecular weight distribution and the smaller molecules in the sample will always degrade by reverse polymerization. The degradation curve obtained experimentally will therefore be a result of the tendency for the elimination of these small molecules leading to an increase in the molecular weight and the partial degradation of the large molecules in the sample leading to a decrease in the molecular weight.

The behaviour of the highest polymer can be used to obtain some further information about the actual mechanism of depolymerization reactions. The osmotic data required for this calculation are as follows:

The values of $(\pi/c)_0$ for polymer at the start was 0.35 and for the polymer after 54 % had been removed as monomer was 0.78. But $0.35 \times \frac{1.00}{1.28} = 0.76$ which is close to the experimental value of 0.78 thus proving that the *number* of polymer molecules remains constant during degradation.

From the above results the number of polymer molecules remains the same even though there has been a considerable diminution in molecular weight. This could be explained in two ways: (a) fission of the polymer occurs at random along the chain, elimination of monomer ensues, but the two polymer radicals combine before complete disintegration of the polymer radicals; (b) fission of the polymer occurs at the ends exclusively, elimination of monomer occurring from the single radical produced. Polymer radicals are stabilized by a disproportionation reaction. Evidence will be presented below to support this latter mechanism.

Figure 2 shows that the polymer sample which was prepared using very pure monomer has exactly the same degradation characteristics as ordinary polymer of roughly the same molecular weight. It has been suggested by R. W. Hall (1947) that polystyrene degrades first by rapid chain scission to shorter fragments which in turn degrade stepwise to monomer. It seemed possible that these points which were vulnerable to rapid attack might be due to an impurity, most likely oxygen, in the molecule. This is especially so considering the copolymer results which are discussed in a later paper. Although polymethyl methacrylate is much less liable to contain oxygen, it nevertheless appears necessary to make certain that distillation of the monomer under reduced pressure and subsequent polymerization *in vacuo* made no difference to the degradation characteristics. This was found to be the case.

In table 2, which gives the results of the degradation of 179,000 polymer, the temperatures are shown at which the degradations were carried out. The lowest

temperature mentioned, 170° C, was the lowest at which degradation could be detected. Figure 2 shows that all these points lie on the same degradation curve indicating that there is no temperature effect in the reaction. All the other degradations were started at 220° C and if high degrees of degradation were required then the temperature was raised to 260° C when the rate at 220° C became too slow.

TABLE 2. DEGRADATION OF 179,000 POLYMER

% degradation to monomer	mol.wt.	mol.wt. as % of original	temp. (° C)
0	179,000	100	—
4	182,000	101	170
17	182,000	101	220
30	167,000	93	220
36	162,000	90.5	240
49	141,000	79	220
55	124,000	69	260

Effect of molecular weight upon rate of production of monomer

In the last section, molecular weight determinations indicated clearly that the degradation reaction mechanism was essentially reverse polymerization. Although indications were obtained that the initial breaking of the polymer molecule occurred at or near the end of the chain, no definite evidence could be obtained in this direction from these measurements. On the other hand, it appeared possible that a study of the rates of production of monomer from samples of different molecular weights might produce fairly definite information as to the location of these initial breaks.

The two extreme possibilities for initial breaking appear to be: first, random breaking along the length of the chain, each bond in every molecule having the same probability of rupture; and secondly, exclusive breaking at the ends of the chains, each chain having the same possibility of rupture.

If initial rupture were random then the probability of a molecule being broken would be proportional to the length of that molecule, so that the large molecules would be eliminated from the sample more quickly than the small ones. In consequence, a small decrease in molecular weight would be expected during degradation. This decrease would necessarily be small and possibly quite undetectable, so that molecular weight measurements could not be expected to give a direct answer to the problem. On the other hand, if initial rupture took place at the ends of the chains, each chain would have the same probability of degradation and so there would be no change in distribution and therefore no change in molecular weight due to the breaking of one size of molecule rather than another.

Rate measurements, on the other hand, should show up the difference between the two possibilities much more clearly.

If the breaks occur in the body of the molecule and each bond has the same probability of rupture then the number of initial breaks which occur in any sample of polymer, no matter what the molecular weight may be, will be the same. But, for each break which occurs one whole molecule will be eliminated. Therefore, since the rate of initial breaking in all samples will be the same, the rate of production of monomer will be proportional to the size of the molecules.

If, however, the ends of the molecules are the vulnerable points the rate of initial rupture will be proportional to the number of ends, that is, inversely proportional to the molecular weight. Again, however, the number of molecules produced per break will be proportional to the length of the molecule so that the overall rate of production of monomer, which is the rate of initial breaking multiplied by the number of monomers produced per break, will be quite independent of molecular weight except when the value is high.

For any given sample of polymer, complete rate of degradation against percentage degradation to monomer curves were obtained in the following way. A sample of polymer was placed in the heating tray and heating was started with the thermostat set at the required temperature. After the volatiles had been eliminated (see preceding paper) a reading of the Pirani gauge voltage (V) was taken and the heating stopped. As the polymer cooled and degradation slowed down, V decreased, until it reached a steady value which was V_0 , when degradation had stopped altogether. The value of $(V^2 - V_0^2)/V_0^2$ could then be obtained and converted to a rate in g./sec. by means of the curve in figure 4 of the preceding paper. The monomer collected in the liquid air trap was distilled into the calibrated capillary to give the extent of degradation. Heating was then restarted and the whole process repeated so that only one sample of polymer was required to obtain a complete curve.

For comparison purposes the rates of degradation were all measured at 220° C.

Figure 3 shows a series of results obtained in this way.

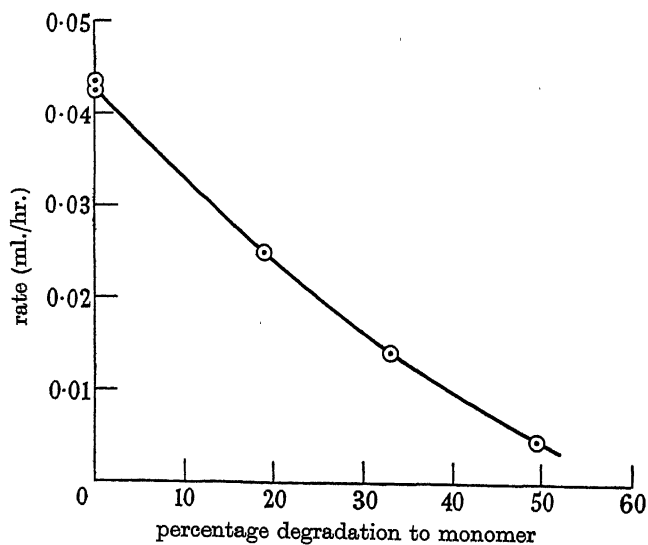


FIGURE 3. The effect of degradation on the rate of production of monomer from benzoyl peroxide catalyzed polymethyl methacrylate having a molecular weight of 36,000.

Owing to the presence of the volatiles no direct measurement of the initial rate could be made. When the whole curve had been obtained as described above it was a simple matter to extrapolate back to zero degradation and obtain the initial rate. The values obtained by this method and that described in the previous paper were 0.0425 and 0.0435 ml./hr. respectively.

Complete rate-percentage degradation curves were obtained for a number of polymer samples and all were found to have the same general structure. It was therefore decided that the initial rates would give the best measure of the relative rates of degradation of a series of samples.

All of these samples had been prepared by benzoyl peroxide catalysis and therefore had the same structure. The results are shown in table 3.

TABLE 3. INITIAL RATES OF DEGRADATION OF 0.2 G. SAMPLES OF BENZOYL PEROXIDE CATALYZED POLYMERS

M_n	$\frac{V^2 - V_0^2}{V_0^2}$	rate (ml./hr.)	% polymer decomposed/sec.
36,000	0.97	0.0425	0.00590
71,000*	1.01	0.0445	0.00618
122,000*	0.99	0.0435	0.00604
190,000*	1.03	0.0455	0.00632
356,000*	0.90	0.0390	0.00542
500,000	0.75	0.0320	0.00444
94,000	0.98	0.0430	0.00597
179,000	1.02	0.0450	0.00625
725,000	0.62	0.0265	0.00368

* Interpolated from viscosity molecular weights.

From these results it is obvious that the rate of degradation of polymer samples of the same general structure is constant for molecular weights between 36,000 and 200,000. These results give direct evidence that degradation is initiated at the ends of the chains. Above 200,000 the rate appears to fall away gradually but this is not surprising since molecular weight measurements have already indicated that the higher molecular weight molecules do not disintegrate completely but are prematurely terminated in the way already indicated.

REFERENCES

- Blatz, P. J. & Tobolsky, A. V. 1945 *J. Phys. Chem.* **49**, 77.
 Fuoss, R. M. & Mead, D. J. 1943 *J. Phys. Chem.* **47**, 59.
 Hall, R. H. 1947 *Faraday Soc. Disc.*, *The labile molecule*, p. 396.
 Kuhn, W. 1930 *Ber. dtsh. chem. Ges.* **63**, 1503.
 Masson, C. R. & Melville, H. W. 1946 *Nature*, **157**, 74.
 Masson, C. R. & Melville, H. W. 1949 *Proc. Instn Rubber Ind.* (in the Press).
 Simka, R. 1941 *J. Appl. Phys.* **12**, 569.

The thermal degradation of polyvinyl compounds

III. The effect of inhibitors and end groups on the degradation of polymethyl methacrylate

BY N. GRASSIE AND H. W. MELVILLE, F.R.S.

Chemistry Department, University of Aberdeen

(Received 3 March 1949)

An investigation has been made of the effect of various end groups upon the degradation of polymethyl methacrylate chains. Inhibition of the degradation reaction by 1, 4 diamino anthraquinone has also been studied.

That small amounts of such a substance do inhibit the reaction shows conclusively that the mechanism is a chain reaction involving free radicals. 1, 4 diamino anthraquinone and its leuco compound have equal inhibitory power indicating that the dyestuff is the active agent. Experiments show that if leuco compound is used, it is very quickly converted to the dyestuff. These experiments, combined with viscosity measurements, gave additional evidence that in degradation the molecules break initially at the ends.

Comparison of the rates of degradation of benzoyl peroxide catalyzed polymers, photopolymer and polymer initiated and terminated by diphenyl cyanomethyl radicals show that double and single bond ends, formed by disproportionation, break most easily. The energies of activation of degradation through them are 30,000 and 42,000 although it is difficult to decide which is which. The energy of activation for degradation through diphenyl cyanomethyl end group is 47,000. The energetics are discussed.

INTRODUCTION

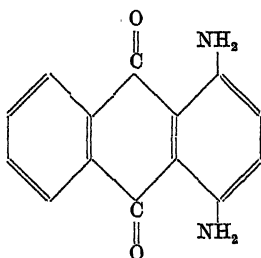
In part II it was shown that polymethyl methacrylate degrades by first being rendered unstable at the ends of the polymer chains, this being followed by a chain degradation reaction yielding only monomer. This instability which is produced at the ends of the molecules is most likely due to the formation of a small free radical and also a large free radical. The rapid stripping off of monomer units from the large radical is a reasonable assumption since it is known that large radicals are in fact thermally unstable. Free radical reactions, however, are susceptible to retardation by traces of inhibitors which react with and destroy the free radicals and so stop the radical chain reaction. If such an inhibitor could be found in this case then it would yield additional evidence that the reaction proceeds by a free radical mechanism.

CHOICE OF AN INHIBITOR

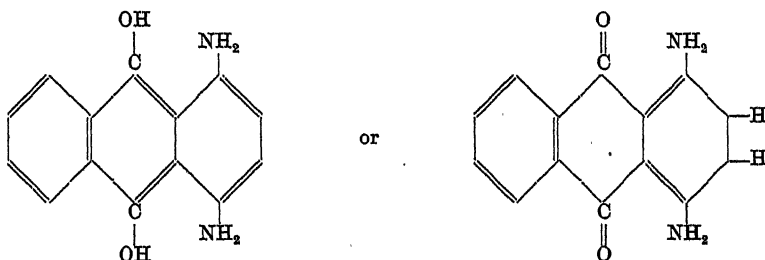
All the usual inhibitors are useless for this purpose owing to the unusual conditions under which the reaction is carried out. The first essential is that the substance must be stable chemically at 220° C and preferably at even higher temperatures. Secondly, the substance must be involatile at 220° C, even at a pressure of only 10^{-6} mm. Finally the inhibitor chosen must be soluble in the polymer so that there may be good contact between polymer and inhibitor molecules. Hydroquinone is a common inhibitor for polymerization reactions, but it would be eliminated very quickly from

the heating tray if the pressure in the still were decreased to 10^{-6} mm. even at ordinary temperatures.

There are certain vat dyes of the anthraquinone type which are used to colour polymethyl methacrylate and which are soluble in it. These dyes may be volatilized in air without decomposition. It was thought possible that the leuco compound of one or other of these substances might act as an inhibitor by splitting off its leuco hydrogen atoms which would act then as chain breakers. On the other hand the dyestuff itself might inhibit the reaction by reacting with the degrading chains in much the same way as it reacts with hydrogen atoms to form the leuco compound. The dyestuff finally chosen was 1,4 diamino anthraquinone which is a maroon coloured substance.



The leuco compound is brown and there appears to be a difference of opinion as to whether the leuco hydrogen atoms are attached in the normal way to the quinone oxygen atoms or in the 2, 3 position to the two amino groups.



For the present purpose a definite choice between the two structures is of no importance.

It remained to be shown whether the compound was involatile under the conditions of temperature and pressure used in the degradation experiments.

It was found possible to maintain the dyestuff in the tray of the still at 220°C and under a pressure of 10^{-6} mm. for an hour without any appreciable evaporation taking place. On slowly raising the temperature still further nothing happened until a temperature of about 240 to 250°C was reached when the dyestuff evaporated quickly and coloured the inside of the glass of the still.

The same kind of experiment was carried out with the leuco compound, when it was found possible to maintain it at 280°C for an hour without appreciable evaporation. On raising the temperature much above 300°C , however, evaporation did occur and the inside of the still was coloured brown. The dyestuff may then be used

at temperatures below 240°C and the leuco compound at temperatures below 300°C .

Another important point is the fact that the leuco compound is not decomposed at this temperature in absence of oxygen. It is very unstable in presence of oxygen and gives up its leuco hydrogen atoms forming the dyestuff, even on gentle heating in air or at ordinary temperatures when dissolved in benzene.

INHIBITION AT 220°C

The molecular weight of the dyestuff is 238, and that of the polymer used for these experiments was 94,000. In order to test first of all whether the dyestuff or the leuco compound are inhibitors for the reaction roughly equimolecular proportions of dye and polymer were used so that enough dye might be present to terminate each chain at least once. Two samples of polymer, each 0.4 g., were weighed accurately into weighing bottles and 0.001 g. of dyestuff or leuco compound was added, one to each sample of polymer. 0.2 g. of each of these mixtures was heated in turn in the still to 220°C and the curves obtained on plotting Pirani gauge volts against time are shown in figure 1. For comparison there is a curve obtained from a similar experiment with 0.2 g. of pure polymer without any addition of dye or leuco compound.

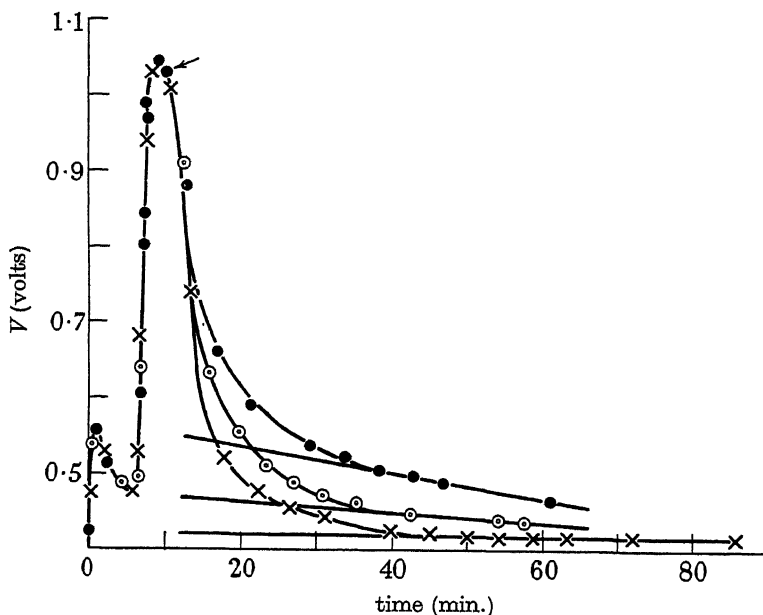


FIGURE 1. Inhibition of the degradation reaction of polymethacrylate of molecular weight 94,000 by 1,4 diamino anthraquinone. ● polymer; × polymer + 1 mole dye or leuco dye; ○ polymer + 0.25 mole dye or leuco dye.

The two maxima due to elimination of volatiles are very accurately reproduced in each case, but after this the rate falls practically to zero both with dye and leuco compound present, indicating that the reaction is almost completely suppressed by both of these compounds. The actual values of initial rate are shown in table 1.

The third result shown in table 1 was obtained with only 0.25 mole of dye per mole of polymer. It will be seen that the rate of evolution of monomer falls to about half its normal value.

That small amounts of foreign substances do inhibit the reaction so markedly, is conclusive proof that the reaction proceeds by a chain mechanism. It appears too that roughly one dye molecule is required per polymer molecule to stop the reaction almost completely. This point will be discussed later.

The fact that the dyestuff is as efficient an inhibitor as the leuco compound seems to indicate that the dyestuff is the active inhibitor.

TABLE 1. . RATES OF INHIBITED REACTION

	% polymer decomp./sec.
pure polymer	0.00618
polymer + 1 mole leuco compound or dye	0.00055
polymer + 0.25 mole dye	0.00257

Normally, when a reaction exhibits an induction period, that induction period will end when all the material causing it is used up. The length of the induction period may sometimes be used to gain information about the nature of the reaction and so in the experiment where equimolecular amounts of dye and polymer were used heating was continued for 6 hr. in an attempt to detect some increase in the rate of evolution of monomer. No increase in rate could be found. Considering the amount of dye present and the normal rate of evolution of monomer without dye, all the dyestuff must have been used up quite quickly if one dye molecule was prematurely stopping the degradation of each polymer molecule. The only explanation appears to be that not only does the dyestuff stop the polymer radical degrading but once it has reacted with the radical it also makes it difficult for degradation to start again. According to this theory, if there are enough dye molecules present to react with each degradable chain end only once, then that polymer sample will be completely prevented from degrading.

Each dye molecule is theoretically capable of stopping two chains, because there are two points in the molecule which can add on a hydrogen atom. On the other hand, each polymer molecule has two points through which it might degrade, so that an equimolecular amount of dye would just be the critical quantity to stop all degradation. If this were the case then the polymer molecule might finally incorporate dyestuff molecules near to the terminal groups.

Further information about the mechanism of the action of the dyestuff can be furnished from a study of the trend of molecular weights during the course of the inhibition. It is unlikely that the dye could interfere with the initial fission of the polymer radicals at one or other end. It is more probable that the polymer radical interacts with the dye to form a more stable radical which cannot easily eliminate monomer. This radical then is stabilized by interaction with another similar radical. This interaction may once again involve either combination or disproportionation. In the former case the molecular weight of the polymer would double; in the latter

it would remain practically the same, except for the incorporation of the dye molecule at the end of the chain. The precise point of attack on the dye molecule is as difficult to decide as the location of the hydrogen atoms in the leuco dye.

It would be difficult to measure osmotic molecular weights of polymer samples containing dye molecules because even small amounts of these molecules might upset the osmotic pressure readings, making them abnormally high; alternatively they might diffuse through semipermeable membranes. For this reason viscosimetric measurements were made to see whether there was a large rise or fall in viscosity which would indicate a change in molecular weight and thus provide evidence for one or other of the mechanisms mentioned above.

It is a general observation that $\eta_{sp.}/c$ plotted against $\eta_{sp.}$ gives a linear plot for low concentrations of high polymer solutions ($\eta_{sp.}$ = specific viscosity, c = concentration). $\eta_{sp.}$ must tend to zero as c tends to zero, and so the intrinsic viscosity, $\frac{\eta_{sp.}}{c}$, of a solution may easily be obtained by measuring the specific viscosity at $c \rightarrow 0$

various concentrations, plotting $\eta_{sp.}/c$ against $\eta_{sp.}$, and extrapolating to zero specific viscosity. The intrinsic viscosities obtained in this way were used to determine whether the average size of the molecules had changed after the degradation reaction with dye present.

These experiments were carried out with the 94,000 polymer. With this polymer no fall in molecular weight occurs until at least 30 % of it has been degraded. Degradation was not therefore allowed to proceed beyond this point.

For these experiments, equimolecular amounts of dye and polymer were used. Owing to the extreme slowness of the reaction at 220° C with dye present, heating was continued for about 30 hr. to degrade a sample to the extent of 27 %.

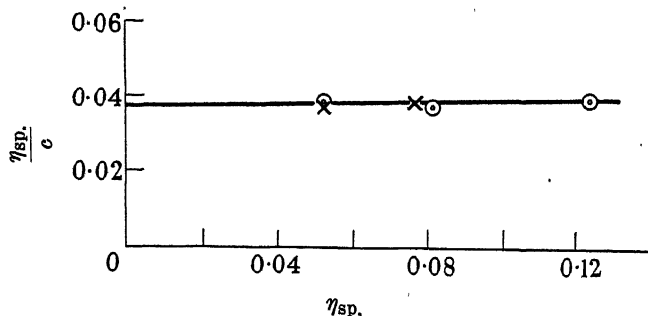


FIGURE 2. Viscosities of 94,000 polymer before and after degradation in presence of an equimolecular amount of 1,4 diamino anthraquinone. $\odot$ original polymer; $\times$ 27 % degraded polymer.

The results obtained are shown in figure 2 and table 2. These results show that there is no difference between the viscosity of a solution of the original polymer and a solution of the polymer degraded to the extent of 27 % in presence of dye. This again is in agreement with the fact that the molecule breaks initially at the end. Since there is no rise in viscosity it appears that each dye molecule can only stop the

degradation of one polymer molecule, and, linking this with the fact that equimolecular amounts of dye are sufficient completely to inhibit the reaction, it would appear also that the polymer can only degrade through one of its ends.

TABLE 2. VISCOSITY MEASUREMENTS

benzene time (sec.)	polymer time (sec.)	$\eta_{sp.}$	conc. (g./1000)	$\eta_{sp.}/c$	$\eta_{sp.}/c_0$
(a) pure undegraded 94,000 polymer					
74.7	83.9	0.1232	3.12	0.0395	} 0.037
74.7	78.6	0.0522	1.36	0.0384	
(b) 94,000 polymer, 27 % degraded at 220° C with an equimolecular amount of dye present					
73.8	79.5	0.0770	2.01	0.0384	} 0.0365
73.8	77.7	0.0527	1.41	0.0375	

THE REACTION AT 280° C

It has been shown that the dyestuff and its leuco compound have approximately the same power of inhibiting the reaction and hence it seems certain that the dyestuff is the active substance. It therefore appears that the leuco compound must be converted to the dyestuff very early in the reaction.

It was shown that the dye is appreciably volatile at about 240° C, whereas the leuco compound is involatile at 300° C. If experiments were to be carried out at 280° C it might be possible to detect just when leuco compound was converted to dye, provided the dye is not too tightly bound to the polymer to be volatilized at this temperature.

As a preliminary, a sample of polymer and dyestuff were heated quickly to 220° C. After that the heating current was cut down so that the temperature of the polymer rose only very slowly. Up to 240° C there was still complete inhibition, but above 240° C, the dyestuff volatilized. At the same time the rate of the degradation reaction increased markedly.

This experiment showed that the dye was still volatile in presence of polymer and so it should be practicable to follow the conversion of leucodye to dye, in presence of polymer, at temperatures above 240° C.

Once again an equimolecular mixture of 94,000 polymer and leuco dye was placed in the heating tray. After evacuating the still, the regulator was set at 280° C and the heating current switched on. The curve obtained on plotting Pirani gauge volts against time is shown in figure 3. If the rise in temperature had been checked below 250° C the curve would have continued to fall steeply at *A*, and would have flattened out just a little above 0.38 which was zero in this case. At *A*, however, there was a distinct reddening of the inside of the still and the rate of the reaction increased to a value comparable with that which would have been obtained if no dye had been present. This experiment showed conclusively that the leuco compound is converted to dye very early in the reaction.

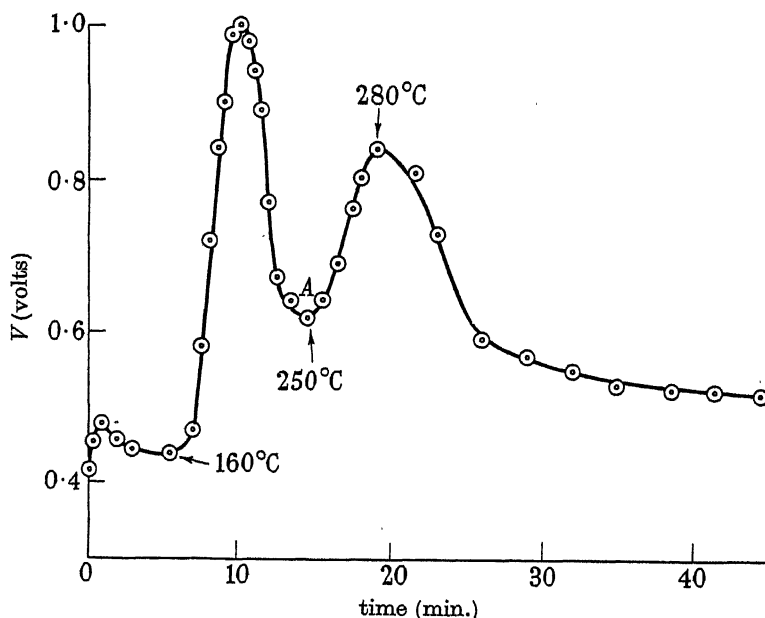
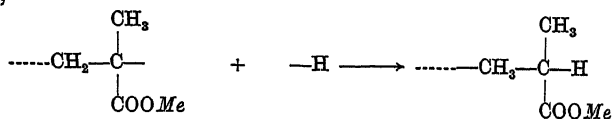


FIGURE 3. The degradation reaction of 94,000 polymer in the presence of an equimolecular amount of leuco 1, 4 diamino anthraquinone at temperatures up to 280° C. Zero = 0.38.

REACTION OF LEUCO COMPOUND WITH MONOMER AND POLYMER

Since leuco compound cannot exist in contact with degrading polymer, it was necessary to find how the hydrogen atoms of the leuco compound were removed.

In a sample of degrading polymer, three types of molecule are present, namely, monomer, dead polymer, and long chain free radicals of partially degraded polymer. If the leuco hydrogen atoms could react with a degrading chain making it stable by simple addition,



this reaction would be very difficult to detect, since these chains would have the same end groupings as the original chains and they would degrade in the same way. These hydrogen atoms would be very quickly used up, and soon all inhibition would be proceeding by the blocking of the ends of the chains by the large dye molecules.

If neither dead polymer nor monomer can react with leuco compound, however, the conversion of leuco compound to dye must proceed by some such mechanism as indicated above. To find whether in fact either of these could react with leuco compound the following experiments were carried out.

Three glass tubes were filled and sealed *in vacuo* after having been thoroughly degassed by alternate freezing in liquid air and thawing.

The first tube acted as a blank and contained only 10 ml. of toluene together with 0.03 g. of leuco dye. The second tube contained similar amounts of toluene and leuco dye together with 1 g. of polymer. The third tube contained the same amounts of

toluene and leuco dye but instead of polymer a few drops of monomeric methyl methacrylate were added. The liquid in these tubes, as prepared, was coloured brown by the leuco dye, and even after remaining at room temperature for 3 days no change in colour took place in any of them.

All three tubes were then heated to 110° C, in boiling toluene, for 30 min., when marked changes took place in the tubes containing polymer and monomer. There was still no colour change in the tube containing only toluene. The tube containing polymer had changed to a reddish colour soon after heating was begun but did not darken further on prolonged heating. On the other hand, the tube containing monomer changed to a deep red colour which became very intense as heating proceeded.

The hydrogen atoms of the leuco dye therefore react with monomer at an elevated temperature. The slight coloration produced with polymer was probably due to either traces of monomer or the possible reaction of the leuco hydrogen atoms with traces of acetone in the polymer. Another possibility arises from the fact that half the polymer chains are terminated by double bonds which may react in much the same way as the double bond in monomer. That it might be due to a reaction with free radicals produced from slow degradation at this temperature is a remote possibility, because the colour developed soon after heating was started and did not darken perceptibly on prolonged heating.

EFFECT OF END GROUPS ON THE DEGRADATION REACTION

It has been established by rate measurements in part II, and by dyestuff experiments earlier in this paper, that the degradation reaction of pure homogeneous methyl methacrylate polymers is initiated at the ends of the chains. In this and the previous papers all the polymer molecules investigated differed only in chain length, the chain ends in every case having exactly the same structure. Any differences in the nature of the reaction of different polymers has been due solely to differences in molecular weight.

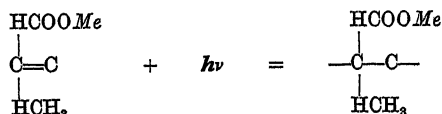
A further proof of terminal fission might be provided by altering the nature of the end groups, for it is certain that the precise structure of the terminal groups will determine the ease with which they are eliminated from the polymer molecule. By a study of the effect of the introduction of different end groupings it may be possible to gain yet more insight into the nature of the reaction.

The molecules already studied have a terminal structure consisting of either a benzoate or phenyl group at one end, and a double or single bond produced by disproportionation in the growing polymer, at the other. It is not yet known through which of these ends the molecules degrade more readily, or whether any of them tend to block the ends of the chains completely.

To attempt to answer such questions, degradation of samples of two other polymers have been carried out. These are (a) a photopolymer, and (b) a polymer having diphenyl cyanomethyl groups at each end. The preparation of the polymers is described below. In the first of these, both ends are terminated by either double or single bonds, and in the second, both ends are terminated by massive aromatic groups which might have a similar effect upon the chain end as benzoyl or phenyl groups. In this way it was possible to study separately, each end of the benzoyl peroxide catalyzed samples.

PREPARATION AND STRUCTURE OF POLYMERS

If perfectly pure monomeric methyl methacrylate is subjected to the action of ultra-violet light it may be made to polymerize. The available evidence indicates that the ultra-violet light activates monomer molecules which then grow as diradicals



Polymerization will take place in two directions and therefore each chain must be terminated twice to become a stable molecule. It is also believed that termination occurs by disproportionation and therefore each molecule will have on the average one double bond and one single bond end. The important point is that both ends consist of pure methyl methacrylate instead of only one end as in the peroxide catalyzed polymer samples. Thanks are due to Dr M. F. Mackay for these samples.

A sample of polymer was prepared by the use of the diphenyl cyanomethyl radical as initiator. At high enough concentrations of this radical the kinetics indicated that not only does the radical start polymerization, it also stops it. Thus the polymer consists of a sandwich of methyl methacrylate units between diphenyl cyanomethyl groups.

DEGRADATION RATES

Comparison of initial rates

It was shown in the previous paper that the molecular weight of a polymer sample has no effect upon the rate of degradation except at high molecular weights where there is a slight lowering in rate. Any change in rate in a sample must therefore be due entirely to the change in initiation rate which in turn will depend upon the nature of the chain ends.

The initial degradation rates which were obtained with photopolymer and diphenyl cyanomethyl polymer are compared in table 3, with the initial degradation rate of benzoyl peroxide catalyzed polymer.

TABLE 3. DEGRADATION OF 0.2 G. POLYMER AT 220° C

polymer	rate (ml./hr.)	% polymer decomp./sec
Bz ₂ O ₂ catalyzed	0.0440	0.00611
photopolymer	0.0650	0.00902
	(0.1 g. = 0.0325)	
diphenyl cyanomethyl	0.0025	0.00035

These results show that an increase in the number of ends formed by disproportionation increases the rate. In photopolymer there are twice as many double and single bond ends and considering the fact that diphenyl cyanomethyl groups strongly diminish the rate of degradation at 220° C (and benzoyl and phenyl may do the same) it might be expected that all degradation of benzoyl peroxide catalyzed polymer

takes place through the double and single bond ends, and hence that the photopolymer degradation rate would be twice the rate of degradation of benzoyl peroxide catalyzed polymer. The photopolymer, however, has a molecular weight of about 500,000 and from part II it would be expected that benzoyl peroxide initiated polymer of this molecular weight would have a degradation rate of about half that of the figure in table 3, namely 0.0045. From the results then it may be deduced that in benzoyl peroxide catalyzed polymer all degradation takes place through the double and single bond ends formed by disproportionation while the benzoyl and phenyl ends resist initiation.

From a consideration of these results it appears most likely that even the double and single bond ends will show a marked difference in ease of initiation. However, it is impossible to obtain either of these two types of end grouping without the other in any polymer sample, so that experiments of the type just described could not give any information on this point.

Investigation of complete rate curves

A typical rate-percentage degradation curve is shown in figure 3, part II, where it may be seen that the rate of production of monomer falls much quicker than would be expected if it were due simply to a decrease in the amount of polymer present. It has already been suggested that the double and single bond ends of the molecules react at different rates and the rapid fall in rate which has been observed may easily be explained by this hypothesis. One of these types of end might be expected to be much more unstable than the other and the initial evolution of monomer will be due mostly to the rupture of the less stable type. As these are used up the rate will fall due to the much slower rate of rupture of the other type of end.

It has already been shown that the initial rate of degradation of photopolymer indicates that degradation takes place through one or other of the types of chain end which have been produced by disproportionation. If, for example, the double bond ends break first, then the initial rate of degradation of photopolymer will indeed be twice that of benzoyl peroxide catalyzed polymer. If the single bonds break next, the rate of degradation of photopolymer should be twice that of benzoyl peroxide catalyzed polymer throughout the whole reaction. If, on the other hand, the benzoyl or phenyl groups are less stable than the single bond ends, then at a later stage in the reaction the rate of degradation of benzoyl peroxide catalyzed polymer should overtake that of photopolymer.

Figure 4 shows the results obtained in the degradation of 36,000 polymer and photopolymer throughout the whole reaction, and it may be seen that the rate of degradation of photopolymer is always about half as large again as that of 36,000 polymer, which is to be expected. Once all these most unstable ends are broken the rate has become too small for accurate measurement and so the temperature was raised to 280° C. Degradation was carried out to an extent of 80 % with both polymer samples, and the figure shows that there is no relative decrease in the rate of degradation of photopolymer but that it continues to degrade throughout the reaction at a rate 1.5 times as fast as that of catalyzed polymer.

As explained above this is sufficient proof that both double and single bond ends degrade, while benzoyl and phenyl ends do not do so at any detectable rate under these conditions.

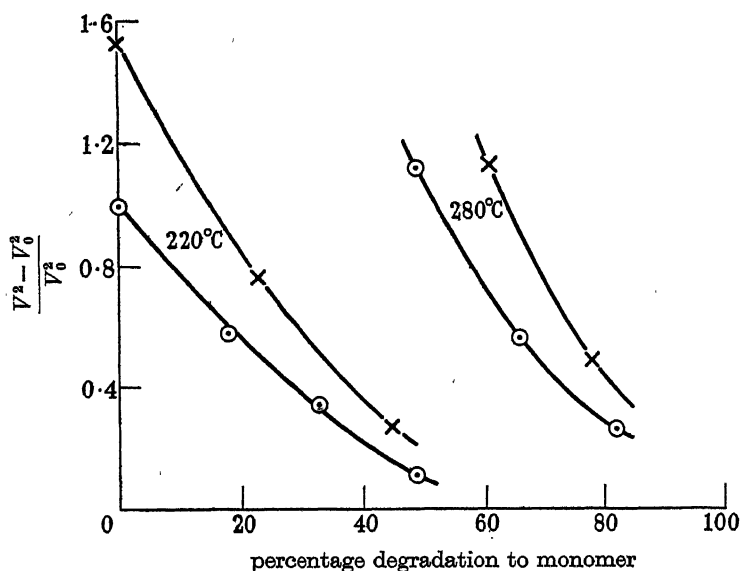


FIGURE 4. Comparison of rates of degradation of photopolymer (○) and 36,000 benzoyl peroxide catalyzed polymer (×).

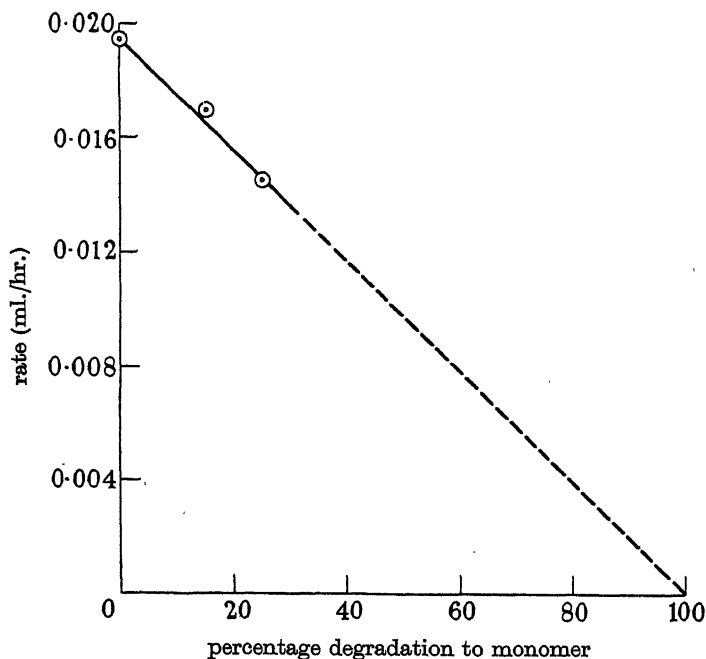


FIGURE 5. Effect of degradation on the rate of production of monomer from diphenyl cyanomethyl terminated polymer.

Whether it is the double or single bond ends which break more readily cannot be discovered by these methods although the problem might be approached from the standpoint of the change in degree of unsaturation of a polymer sample with degradation.

It was shown earlier in this paper that polymer with diphenyl cyanomethyl end groups degraded only very slowly at 200°C. At 280°C, however, the reaction proceeds at a measurable rate. Since the ends of all the chains in such a polymer are identical a fall in rate such as was observed in other polymer samples should not be found. Only a slight decrease should be observed due simply to the decrease in the amount of polymer present.

Results obtained during the degradation of this polymer sample are shown in figure 5.

These results fall in with the general theory since figure 5 shows that there is indeed no rapid fall, the fall in rate up to 25 % at least being due only to the decrease in the amount of polymer present. Degradations were not carried out to a greater extent than 25 % owing to the high temperatures required over relatively long periods of time.

ENERGIES OF ACTIVATION

Experimental determination

Provided the polymer had previously been heated to remove volatiles it was found on reheating it to any fixed temperature that the rate of production of monomer very quickly reached a steady value. From this fact it followed that the value of the energy of activation could be determined in a single run and in a matter of 40 to 60 min.

The polymer was heated quickly to the highest temperature which was to be used in the determination of the energy of activation and as soon as all the volatiles had been driven off (as shown by the rate-time curve) it was allowed to cool. During cooling, by following the decrease in rate on the Pirani gauge, the most convenient temperature for the subsequent determination could be obtained, that is the lowest temperature at which there was a measurable rate of production of monomer.

The temperature regulator was then set at this temperature and the heater current switched on, the readings on the Pirani gauge voltmeter being plotted against time. Readings were taken in the usual manner, that is, as the falling temperature slowly passed the required value. As soon as a constant voltage value had been obtained, the regulator was set 10° C higher and the polymer allowed to come to this new temperature. This was repeated four or five times as shown in figure 6. Once the final value of V had been measured the polymer was allowed to cool and V_0 , the value of V at zero pressure, was obtained.

These results are given in table 4. The values of log rate, measured by $(V^2 - V_0^2)/V_0^2$, are plotted against $1/T$ in figure 7.

The energy of activation corresponds to 31 kcal.

Results

By the method described earlier in this section the energies of activation of all the benzoyl peroxide initiated polymer samples, and the photopolymer, were found

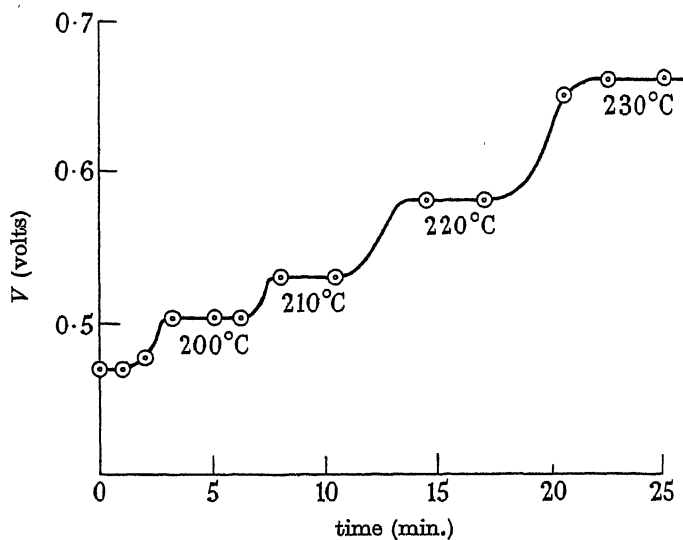
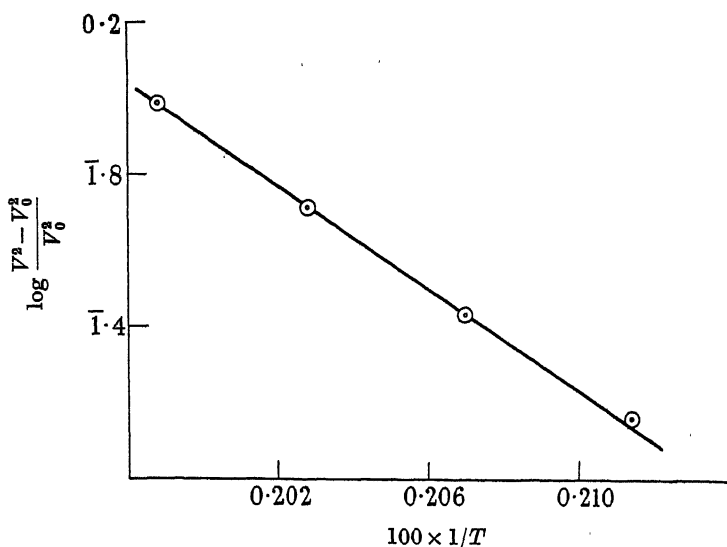
FIGURE 6. Determination of energy of activation I. $V_0 = 0.47$.

FIGURE 7. Determination of energy of activation II.

TABLE 4. DETERMINATION OF ENERGY OF ACTIVATION

$T(^{\circ}\text{K})$	mol.wt. = 179,000. wt. of polymer = 0.1 g.		% polymer decomp./sec.
	$\frac{V^2 - V_0^2}{V_0^2}$	$\log \frac{V^2 - V_0^2}{V_0^2}$	
473	0.1448	1.1608	0.00159
483	0.2715	1.4338	0.00300
493	0.5204	1.7163	0.00574
503	0.9729	1.9881	0.01070

to have values of $31,000 \pm 1,000$ cal. These were measured near the beginning of the degradation reaction, but later measurements indicated a general rise in the energy of activation as the reaction proceeded. Figure 8 shows the result of the measurement of the energy of activation of a typical sample at various stages in the degradation.

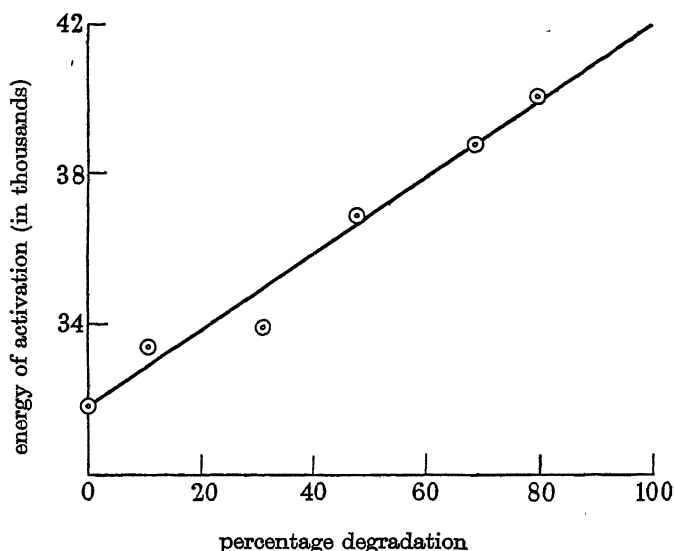


FIGURE 8. Effect of degradation on the energy of activation of benzoyl peroxide catalyzed polymer.

In the reverse polymerization reaction which has already been shown to apply in this case, there are only two steps, these being initiation at the ends of the chains followed by propagation along the chains. The increase in the energy of activation during the reaction must be due to a change in one or other of these steps. In the propagation step, and especially with high molecular weight samples there has been shown to be a definite termination reaction which tends to terminate the chains before they have disintegrated completely. If this were causing the increase in the value of the energy of activation in some way, no increase would be observed during the degradation of 36,000 polymer, which has been shown to be purely reverse polymerization with no premature termination whatsoever. This was shown not to be the case, however, the energy of activation increasing from about 31,000 to 42,000 during the course of the reaction in the same way as with the other polymer samples.

It has already been shown that end groups have a marked effect upon the rate of the reaction, and that degradation proceeds from the ends which have been produced by disproportionation—that is, the double and single bond ends. One of these ends is much less stable than the other, and it seems most probable that the increase in energy of activation during the course of the reaction is due to the fact that as degradation proceeds a bigger proportion of it is taking place through the more stable end which will require a larger amount of energy to eliminate.

It may be deduced that the energies of activation of the degradation reaction through the two types of end formed by disproportionation are 31,000 and 42,000 cal.

In the case of the degradation of diphenyl cyanomethyl end group polymer, since all the ends are of the same nature, it would be expected that the energy of activation would not change during the course of the reaction, in accordance with the fact the rate of monomer production falls proportionately with the weight of polymer remaining in the still. On the other hand, owing to the stability of this polymer, the energy of activation of the degradation reaction would be much larger.

From experiments carried out at temperatures between 270 and 300°C the energy of activation was found to be 47,000 cal. at 2 to 3 % degradation and 47,000 cal. at 25 % degradation. This again supports the theory that the characteristics of the degradation are as already described.

ENERGETICS OF DEGRADATION

Further confirmation of the general chain-like character of the degradation reaction is afforded by an examination of the energetics of the process. If the mechanism suggested is correct this energy of activation for the elimination of methyl methacrylate from the polymer *radical* should be much less than the elimination of the end group from the polymer *molecule*. The former energy of activation is simply the sum of the heat of polymerization (11 kcal.) and the energy of activation (4 kcal.) for the growth of the polymer radical or 15 kcal. This is much less than the lowest value (31 kcal.) for the degradation of the photo polymer. Similarly the low value of 31 kcal. would preclude any suggestion that fission of the polymer molecule could occur at a carbon-carbon bond along the chain. It is not possible to say to which particular bond the 31 kcal. applies until the chemical nature of the radical fragment from the polymer is identified. With the larger polymers the polymer radicals do not completely degrade and hence the overall energy of activation will be composite as occurs in the polymerization reaction itself. Thus

$$E_{app.} = \frac{E_1}{2} + E_2 - \frac{E_3}{2},$$

where E_1 is terminal fission energy, E_2 is the elimination energy for monomer from the polymer radical, and E_3 is the radical-radical interaction energy of activation. E_1 is about 31 kcal., E_2 is 15 kcal. and E_3 is probably zero. Thus in this case $E_{app.} \approx 30.5$ kcal. It is portentous that these should be similar values, but this is in accordance with the fact that the energy of activation is independent of molecular weight even though the *rate* falls off at high molecular weights.

In common with the thermal fission of organic molecules it is to be presumed that the initial fission is a unimolecular reaction in the sense that once the energy has accumulated in the bond to be broken, the actual fission is wholly the concern of that particular molecule.

The temperature-independent factor in the initial fission reaction may be calculated since the rate of evolution of monomer and the length of the polymer radical are known. The data are summarized in table 5.

TABLE 5

nature of polymer	mol.wt. of polymer	fractional loss of wt. of polymer/sec.	energy of acti- vation (kcal.)	temp. (° C)	temp.- independent factor
peroxide	36,000	1.16×10^{-4}	32	220	4.2×10^7
initiated	36,000	4.4×10^{-5}	40	280	6.4×10^7
diphenyl cyanomethyl	—	1.29×10^{-5}	47	280	9.6×10^9

It will be seen that these factors are considerably less than 10^{13} which is the normal factor in unimolecular reactions. So little is known about energy transfers within long-chain molecules that it is not surprising that abnormalities should be observed.

The thermal degradation of polyvinyl compounds

IV. The thermal degradation of the methyl methacrylate copolymers with glycol dimethacrylate and acrylonitrile

BY N. GRASSIE AND H. W. MELVILLE, F.R.S.

Chemistry Department, University of Aberdeen

(Received 3 March 1949)

The effect upon degradation of cross-links, formed by copolymerizing small amounts of glycol dimethacrylate with the methyl methacrylate, and small amounts of acrylonitrile units in the chains has been investigated.

Cross-links make no difference to the degradation characteristics according to rate measurements. Molecular weights could not be measured in this case owing to the insolubility of the polymer.

The methyl methacrylate-acrylonitrile copolymer degrades in quite a different way from polymer composed of homogeneous methyl methacrylate chains. In this case the chain ruptures initially at or near the acrylonitrile units, the molecular weight falling very rapidly in the initial stages of the reaction. One side of the break is terminated by an acrylonitrile unit which cannot degrade but the other side may sooner or later degrade to the next acrylonitrile unit along the chain. The induction period observed is probably due to the time taken to build up an appreciable concentration of such degradable ends.

A quantitative calculation agrees well with this theory.

INTRODUCTION

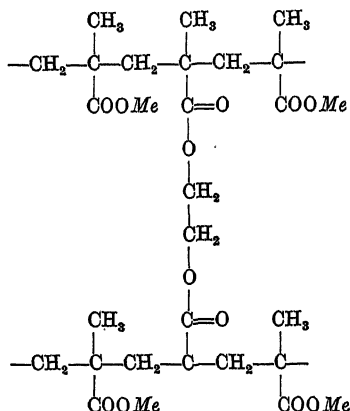
In previous papers the degradation characteristics of pure polymethyl methacrylate chains have been studied and the mechanism of the reaction established. This degradation reaction has been shown to proceed in two well-defined steps. First the chain is initiated by a splitting from the end of the molecule of a small free radical. The remaining polymer radical completely disintegrates with the production of monomer. It has been shown possible to vary the initiation rate by changing the

nature of the end groups in the polymer molecules, although once this initiation has occurred there is no changes in the characteristics of the depolymerization.

The experiments in this paper are concerned with the depolymerization step itself and the effect of changes in the structure of the polymer chain will be described.

A. THE METHYL METHACRYLATE-GLYCOL DIMETHACRYLATE COPOLYMER

The first change was the introduction of cross-links along the chain. This was done by copolymerizing glycol dimethacrylate with methyl methacrylate. The structure at a cross-link would be as follows:



The structure of the main chains has only been slightly changed. Such a modification should not significantly affect the nature of the depolymerization process.

Preparation of samples

In the preparation of these cross-linked polymer samples the method of purification of monomer was exactly the same as that described in part II for the preparation of pure polymethyl methacrylate samples, except that the requisite amount of glycol dimethacrylate was added before purification. The reason for this was that such a small amount of glycol dimethacrylate was used that separate purification by the same method as for the methyl methacrylate would have proved a difficult and wasteful business.

Once the inhibitor had been removed and the mixture of monomer and cross-linking agent dried, purified catalyst was added and the whole poured into a bolt-head flask which was heated in a water bath.

Three samples were prepared under conditions shown in table 1.

TABLE 1. PREPARATION OF CROSS-LINKED POLYMER SAMPLES

sample	temp. (° C)	% catalyst	% G.D.M.
1	100	1	1
2	100	0.25	1
3	100	1	5

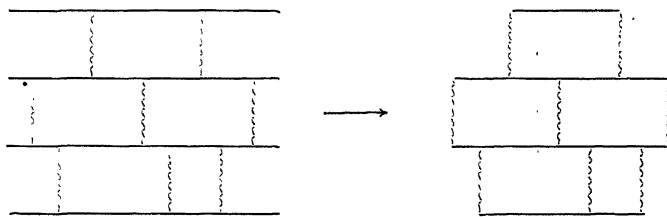
This cross-linked polymer is insoluble in all solvents including monomer, so that as soon as it was formed it precipitated from solution and set to a solid mass in a very

short time. Owing to its insolubility it was quite impossible to obtain it in the form of a fine powder by the method described for pure polymethyl methacrylate.

The semipolymerized mixture was poured on to large watch glasses just before it set solid, and afterwards broken up and roughly ground.

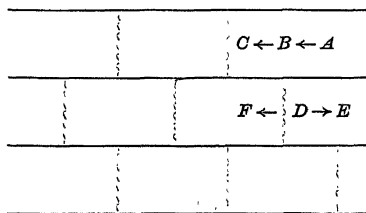
Theoretical

Owing to the insolubility of the polymer it is unfortunate that no information can be obtained about the nature of the sample during the degradation. A number of possible extreme mechanisms for the reaction present themselves from a knowledge of the structure of the polymer. It may be, for example, that even this small change in the structure of the chains themselves will cause degradation to stop entirely when a cross-link is reached. In that case initiation would proceed as before at the ends of the chains, but the reverse polymerization reaction could only proceed as far as the first cross-link in each chain. Very soon all evolution of monomer would stop and a very stable and compact molecule would result.



A second possible extreme mechanism arises from the assumption that a cross-linking unit would make little difference to the degradation. When a chain degradation reaches this cross-link, the unit containing it will be eliminated in the normal way. The glycol dimethacrylate molecule would not, however, be evolved until both chains, of which it has been a part, had depolymerized.

Finally, it appears just possible that while degradation takes place through one vinyl group of a glycol dimethacrylate molecule, the other vinyl group may be affected, becoming unstable, and starting off degradation in the adjacent chain. If, for example, degradation proceeded from *A* to *B*, then at *B* it would branch, and proceed both to *C* and also from *D* to *E* and *F*.



In the first of these cases the rate would be very slow in the initial stages, and very quickly fall to zero. In the second case the rate would not be very much different from that during the degradation of normal benzoyl peroxide catalyzed polymer. It would certainly be of the same order, although the size of the particles or a slight

difference in the reactivity rate of glycol dimethacrylate might cause a slight difference from straight chain polymer.

Lastly, if the third mechanism were correct, then the whole molecule would very quickly disintegrate to monomer at an ever increasing rate.

Experimental and discussion

0.2 g. of each of the three samples were degraded in the heating tray in the usual manner. The same sort of curve was obtained as that shown in figure 9, part I (for ordinary polymethyl methacrylate), except that the second maximum was not nearly so high. This maximum was explained as being due to evolution of solvent or monomer from the inside of the polymer grains, when the polymer melted. Owing to the infusibility of these cross-linked samples, however, these volatiles would not be so easily evolved as with straight chain polymethyl methacrylate, and also the sample would not be so liable to contain volatiles since no solvents were used in its preparation. The values of rate, $(V^2 - V_0^2)/V_0^2$, at 220° C, for different degrees of degradation of sample 3, are given in figure 1, which shows that the degradation proceeds in much the same way as for normal polymer. The initial rates for each

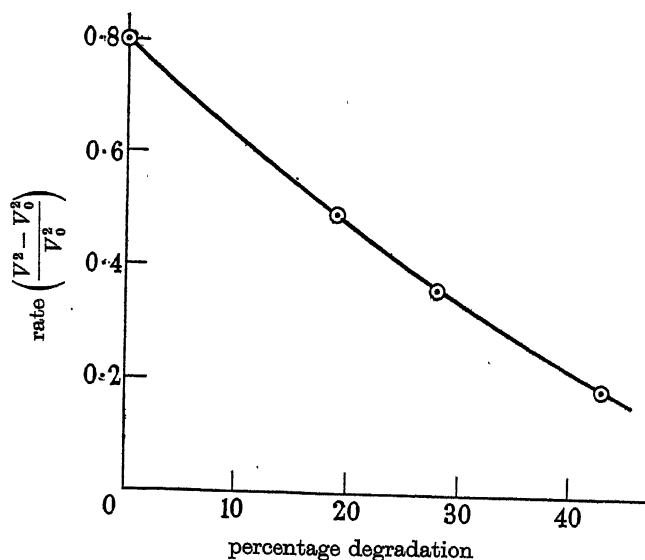


FIGURE 1. Rate of degradation of cross-linked polymer at 220° C.

of the samples are given in table 2, which shows that neither the amount of catalyst present, nor the amount of cross-linking, has any effect upon the rate of degradation. It is clear from these results that the second mechanism suggested is indeed the correct one. The difference between the initial rates of these samples, $(V^2 - V_0^2)/V_0^2 = 0.78$, and straight chain polymethyl methacrylate, $(V^2 - V_0^2)/V_0^2 = 1.0$, is probably due to the relatively large size of the grains and their infusibility, both of which factors make it much more difficult for the monomer to escape from the polymer once it has been formed by degradation.

It has been suggested that part at least of the insolubility of cross-linked polymers is due to the extreme length of the chains which is caused in some way by the cross-

linking agent. It has been shown, however, that in the case of straight chain polymers, the rate of degradation falls away markedly at high molecular weights. But even if the decrease in rate of degradation of these cross-linked polymers was totally due to the length of the chains one would only expect the chains to be about 4000 to 5000 units in length. From this it follows that these polymer samples must be made up of relatively short chains initiated and terminated in the usual way but containing an occasional cross-link which has made no difference to the mean structure of the polymer chains.

TABLE 2. INITIAL RATES OF DEGRADATION OF 0.2 G. SAMPLES OF CROSS-LINKED POLYMERS

sample	$\frac{V^2 - V_0^2}{V_0^2}$	rate (ml./hr.)	% loss of wt. of polymer/sec.
1	0.78	0.0335	0.00439
2	0.75	0.03225	0.00421
3	0.80	0.0345	0.00450
normal peroxide polymer	1.00	0.0440	0.00574

This difficulty of monomer escaping from the polymer may be demonstrated by the following experiment. After all the volatiles had been removed from 0.2 g. of sample 1 it was cooled to room temperature. The regulator was then set to 250° C and heating started once more. The form of the curve obtained on plotting Pirani gauge volts against time is shown in figure 2. When the temperature first reached 250° C the evolution of monomer was extremely slow and only built up to a maximum after about 15 min. Under these conditions it is likely that monomer may undergo side reactions such as repolymerization before it has time to escape from the degrading polymer, and so it may be that the discrepancy in rates between cross-linked polymer and straight chain polymer is due, in part at least, to this slow escape of monomer from the polymer.

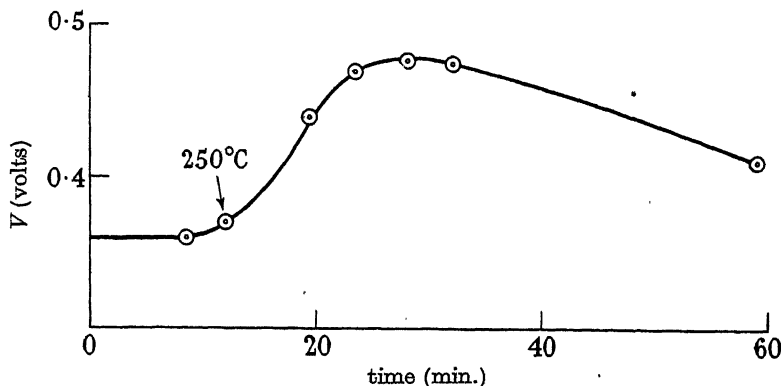


FIGURE 2. Slow attainment of equilibrium in the degradation of cross-linked polymer.

With ordinary polymer the build-up to a steady rate on increasing the temperature takes place almost instantaneously.

B. THE METHYL METHACRYLATE-ACRYLONITRILE COPOLYMER

The first part of this paper was concerned with experiments with a polymer in which a very small change in structure, as far as the degrading chains were concerned, had been introduced. To the degrading chain the change was three atoms away, and as might have been expected it caused little or no change in the degradation characteristics from those of pure polymethyl methacrylate chains.

To take matters a stage further, it appeared desirable to study the effect upon degradation of the introduction into the chain of foreign monomer molecules which had an entirely different structure from methyl methacrylate. It is possible that a methyl methacrylate polymer containing small amounts of the monomer of another degradable polymer such as styrene, would degrade in much the same way as pure polymethyl methacrylate. This, however, is by no means definite since the fact that both methyl methacrylate-methyl methacrylate bonds and styrene-styrene bonds break easily in degradation is not proof that a methyl methacrylate-styrene bond will break as easily. Although such copolymers are worthy of study it would be of more interest at this point to study a copolymer system in which one of the constituents is not degradable and is relatively stable thermally.

Choice of system

In copolymer chemistry it is found that the fraction of each monomer in the copolymer is not the same as the fraction of each monomer in the initial monomer mixture. It is essential, however, in this work to know the exact constitution of the copolymer on which investigations are being carried out.

The elements of a theory of copolymerization were first put forward by Norrish & Brookman (1939). The theory used here is an extension of their work but so far it has only been applied to a very limited number of systems by Alfrey, Merz & Mark (1947), C. C. Price (1946) and Alfrey & Price (1947).

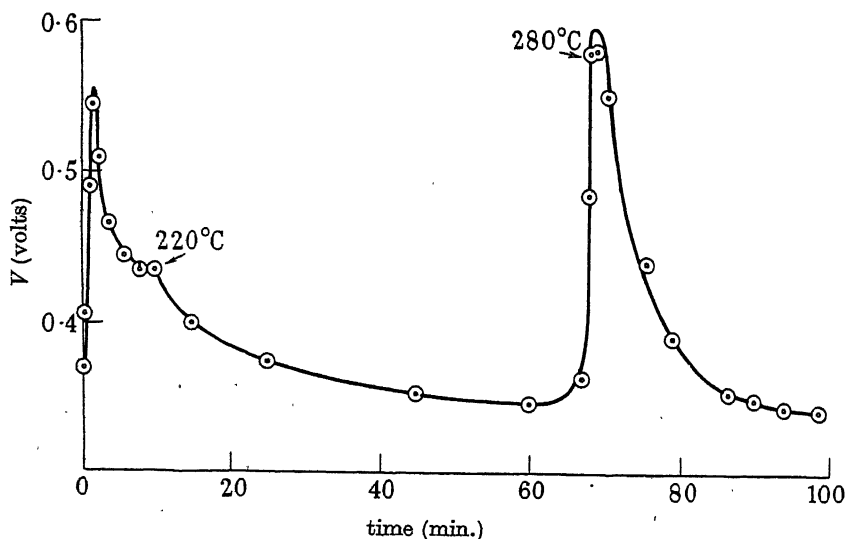


FIGURE 3. The action of heat on polyacrylonitrile.

Since methyl methacrylate was required as the degradable component of the copolymer, the choice of a system was at once very much restricted. The most suitable system appeared to be methyl methacrylate-acrylonitrile since polyacrylonitrile itself is stable thermally to quite high temperatures.

To make quite certain that polyacrylonitrile is thermally stable under all conditions of temperature used during experiments, a sample of this polymer was heated in the still to 220° C and then to 280° C. The curve obtained on plotting the Pirani gauge volts against time is shown in figure 3. Once volatiles have been evolved no further degradation takes place. The temperature was maintained at 280° C for half an hour. The curve has only one maximum which would indicate that only the volatiles adsorbed on the surface of the polymer had been evolved, and therefore that the polymer had not softened or melted. On opening the still this was found to be the case, the polymer still being in the form of a fine powder instead of a solid fused mass as with polymethyl methacrylate.

Copolymerization theory

From the copolymerization theory mentioned above the following equation may be derived, which links the proportion of the monomers in the monomer mixture and in the initial polymer formed from this mixture.

$$\frac{b}{a} = \frac{B}{A} \times \alpha \times \frac{A + \beta B}{A + \alpha B},$$

where b and a are the mole fractions of the monomers in the initial polymer, and B and A are the mole fractions of the monomers in the monomer mixture.

$\alpha = k_{ab}/k_{aa}$ = ratio of rates of addition of A and B monomers to A long chain-free radical ends, and

$\beta = k_{bb}/k_{ba}$ = ratio of rates of addition of A and B monomers to a B long chain-free radical end.

Lewis, Mayo & Hulse (1945) found that the relative rates of addition of monomer to radical, in the formation of the methyl methacrylate-acrylonitrile copolymer, are as shown in table 3.

TABLE 3. RELATIVE RATES OF ADDITION OF MONOMERS
TO LONG CHAIN RADICALS

radical	monomer	relative rates
acrylonitrile	acrylonitrile	1.0
acrylonitrile	methyl methacrylate	7.0
methyl methacrylate	methyl methacrylate	1.0
methyl methacrylate	acrylonitrile	0.83

If A = methyl methacrylate, and B = acrylonitrile, then $\alpha = 0.83$ and $\beta = 0.14$.

Using these values of α and β , figure 4 shows the connexion between mole fractions of monomers in the monomer mixture and in the initial polymer formed from that monomer mixture.

It may be noted, of course, in the preparation of these polymer samples, that the theory only gives information about the initial polymer formed because as the polymerization proceeds the composition of the monomer mixture will change and so affect the composition of the polymer formed from it.

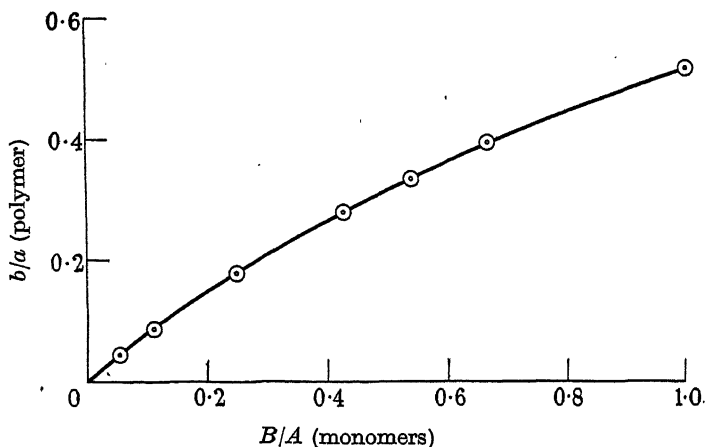


FIGURE 4. Acrylonitrile-methyl methacrylate copolymer composition.

Preparation of samples

The relative monomer-free radical reactivities which were obtained by Lewis *et al.* (1945), were calculated from copolymer samples which were prepared under very closely controlled conditions. Although it is difficult to ensure that the conditions of Lewis *et al.* are maintained during copolymerization thus obtaining similarity in initiating and terminating, the composition of the copolymer is fortunately not affected by these reactions and in fact is completely determined by monomer composition.

The benzoyl peroxide which was used as a catalyst for the polymerizations was purified in exactly the same way as already described for the preparation of benzoyl peroxide catalyzed samples of pure polymethyl methacrylate. Lewis *et al.* used very small catalyst concentrations (0.1 mole %) and hence the catalyst was measured into the reaction vessel as a standard solution in benzene, the benzene being evaporated off by means of a filter pump before adding the monomers.

The acrylonitrile was obtained from Imperial Chemical Industries (Dyestuffs Division). It was carefully fractionated, only a few ml. at the beginning and end of the distillation having to be discarded as not boiling exactly at 77° C.

The monomeric methyl methacrylate was freed from inhibitor by repeated shaking with sodium hydroxide solution. It was finally washed two or three times with distilled water to remove excess sodium hydroxide and dried overnight over calcium chloride. Owing to the possible effect of even trace impurities upon the constitution of the copolymer, it was necessary to purify the monomer further. Lewis *et al.* distilled it under reduced pressure and stored the resultant product in a refrigerator. The roughly purified and dried product was distilled at 35° C under a pressure of 30 to 35 mm. of mercury. The first distillate was distinctly milky and was discarded.

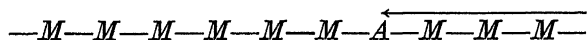
The clear distillate was collected for use and distillation was stopped when about 75 % had been distilled, the residue being discarded. This pure monomer was then stored in a refrigerator until used.

Two samples were prepared with a ratio of monomers such that the initial polymers were made up of methyl methacrylate-acrylonitrile = 410/1 and 40/1. The monomer mixture which was added to the catalyst was thoroughly degassed before being sealed off *in vacuo* and polymerized. Polymerizations were carried out under exactly the same conditions as those of Lewis *et al.* and the polymer was precipitated and powdered as described in an earlier paper. These polymers, which contained only very small amounts of acrylonitrile were found to be soluble in the usual polymethyl methacrylate solvents.

Theoretical

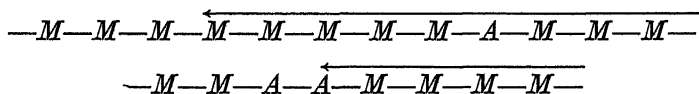
Bearing in mind the degradation characteristics of polymethyl methacrylate and the thermal stability of polyacrylonitrile, various mechanisms may be considered for the degradation of the methyl methacrylate-acrylonitrile copolymer system.

1. Since acrylonitrile itself does not degrade to monomer and is stable thermally, it is possible that the degradation of a polymethyl methacrylate chain may be stopped short by the presence of an acrylonitrile unit.



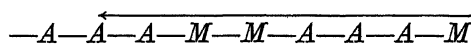
In this case only small amounts of monomer would be evolved and this monomer would represent the average distance from the ends of the chains to the first acrylonitrile unit.

2. On the other hand it may be that when a single acrylonitrile unit is present in the chain degradation will pass right through it. If, however, more than one acrylonitrile units occur together degradation would be stopped since an acrylonitrile-acrylonitrile bond would have to be broken. Only acrylonitrile-methyl methacrylate bonds would have to be broken in the former case.



If this were the mechanism then complete degradation of chains would occur as with pure polymethyl methacrylate as long as the number of acrylonitrile units in the chain was appreciably below 50 % of the total. On approaching 50 % acrylonitrile, the chances of finding two acrylonitrile units adjacent to one another in the molecule would increase.

3. It is just possible that most polymers have the potentiality of degrading although they do not normally do so since initiation cannot occur owing to the nature of the ends of the chains. It has already been shown how, by varying the nature of the end groups of polymethyl methacrylate, the degradation may in some cases be almost completely stopped. If that were the mechanism, then degradation would proceed in all chains which were terminated by methyl methacrylate.



In all three of these mechanisms, degradation, if it takes place at all, does so as in the case of pure polymethyl methacrylate, from the ends of the chains, and in no case would more than one residue be obtained from each partially degraded molecule. From this it follows that the molecular weight of residue against percentage degradation to monomer curves, figure 1, part II, would never fall below the diagonal.

4. On introducing acrylonitrile units into methyl methacrylate chains, acrylonitrile-methyl methacrylate bonds have thereby been formed. Little can be deduced of their properties from the behaviour of methyl methacrylate-methyl methacrylate or acrylonitrile-acrylonitrile bonds, but they may well be unstable and break easily on heating. If this happened, and reverse polymerization took place in the fragments, then the molecular weight would not fall during degradation, provided the two fragments degraded completely. If, however, these degrading radicals were prematurely stabilized then there would be a marked fall in molecular weight and the molecular weight against percentage degradation to monomer curve would lie below the diagonal.

Once again then, as in the case of pure polymethyl methacrylate, it appears possible that rate and molecular weight measurements will give much information as to the mechanism of the reaction.

Rates of degradation

In order to find out, first of all, whether the presence of acrylonitrile units at intervals along the chain has any large effect upon the degradation, an experiment was carried out with 0.2 g. of the 1/410 copolymer sample. Ordinary polymethyl methacrylate would have an initial rate given, at 220°C, by $(V^2 - V_0^2)/V_0^2 = 1.0$ (0.0001148 g./sec.). The curve obtained with a copolymer sample is shown in figure 5, where the initial rate is given by $(V^2 - V_0^2)/V_0^2 = 0.246$ (0.0000282 g./sec.). This falls to $(V^2 - V_0^2)/V_0^2 = 0.101$ (0.0000116 g./sec.) after only 3 % degradation.

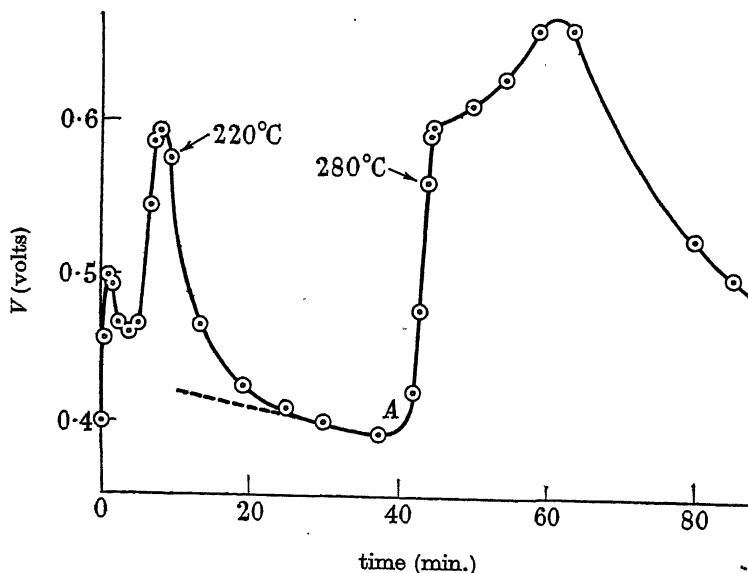


FIGURE 5. Degradation of 1/410 copolymer at 220 and 280°C. Zero = 0.371.

These results indicate that the rate of degradation is cut down markedly by the presence of small amounts of acrylonitrile in the copolymer. At point *A*, in figure 5, the temperature was allowed to rise to 280° C, in order to find out whether at this temperature too the rate was very slow. The rate increased as shown as the temperature rose, but instead of falling away gradually once the temperature had reached 280° C, it remained constant for a short time and then continued to rise, later falling away. The portion of the curve obtained at 280° C appeared to indicate some sort of induction period, although it was not at all well defined. It was decided to search for a similar induction period at 220° C, because at that temperature the rate of production of monomer was much slower, and any increase in the rate would probably be much more sharply defined so that the induction period in turn would be more sharply defined.

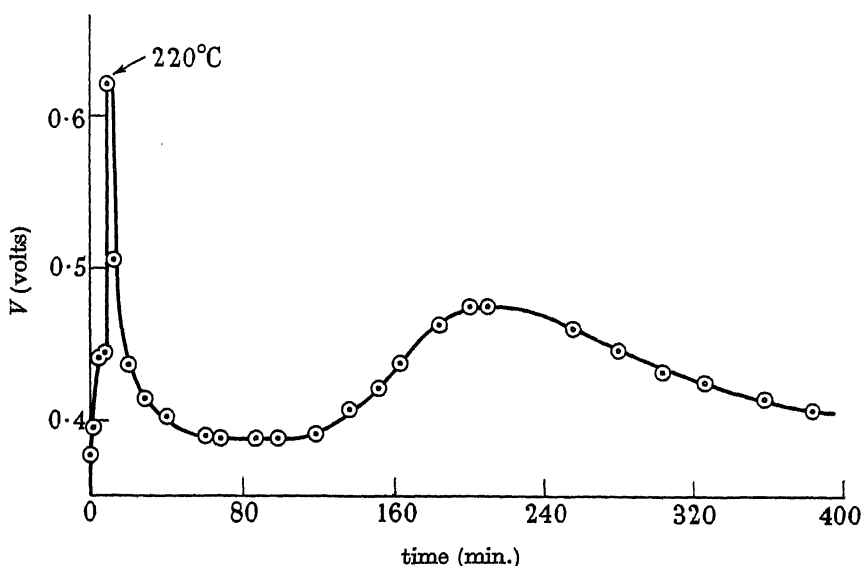


FIGURE 6. Continued degradation of 1/410 copolymer at 220° C. Zero = 0.365.

Figure 6 shows the result of the continued heating at 220° C of 0.3 g. of the same copolymer sample. It will be seen that after about 120 min. the rate did indeed start to rise and reached a maximum 210 min. after the start of the experiment.

The induction period

The induction period obtained at 220° C with 1/410 copolymer appeared to be quite sharply defined, and it seemed that information might be gained as to the role of the acrylonitrile in the reaction by a study of the induction periods of various copolymer samples.

With this end in view, 0.3 g. of the 1/40 sample was heated to 220° C in the still and the result is shown in figure 7. In this case the induction period was not nearly so well defined although it appeared to be of about the same duration as in the previous case. The maximum rate attained was small.

If the induction period had been due to an elimination of acrylonitrile or something of that sort, then it would have been expected to be much longer for the 1/40 sample, although the maximum rate attained would have been more or less the same. It is therefore probably not a true induction period in the sense that during it, something is being eliminated, but rather is it a general slowing down of the rate, caused in some way by the presence of the acrylonitrile units in the molecule.

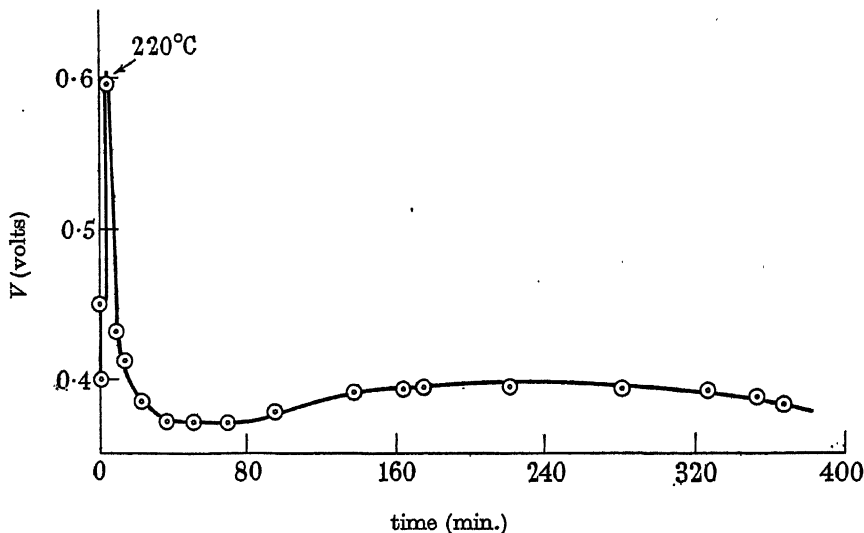


FIGURE 7. Continued degradation of 1/40 copolymer at 220° C. Zero = 0.3575.

Effect of degradation upon molecular weight

Taken by themselves, rate measurements and a study of the induction period could give no more information of the nature of the reaction. The other line of attack, which had been used in the investigation of the degradation of ordinary polymethyl methacrylate, was the measurement of molecular weights of partially degraded samples, and it seemed definite that such experiments might indicate what was happening to the molecules during the apparent induction period—some change must have been taking place which was making them more easily degradable.

Molecular weight measurements were made on degraded samples of both 1/410 and 1/40 copolymers. The results are plotted in figures 8 and 9 and corrected for volatiles in tables 4 and 5.

These two curves indicate that whatever else may happen, two stable involatile fragments are formed when each break occurs in the chain. The only difference in structure between these polymers and pure polymethyl methacrylate resides in the acrylonitrile units which are dispersed along the chain. The breaks must therefore occur at or near the acrylonitrile units since it is already known from previous work that the homogeneous methyl methacrylate parts of the chains do not break in this way.

The limits of molecular weight to which these molecular weight against percentage degradation to monomer curves tend is also of interest. There are on the average,

410 methyl methacrylate units, that is 41,000 in weight, between two acrylonitrile units in the polymer referred to in figure 8 and it may be seen that the molecular weight tends to a value somewhat less than 50,000. In figure 9 the limiting value will be a good deal lower and in this case the weight between acrylonitrile units is only 4000.

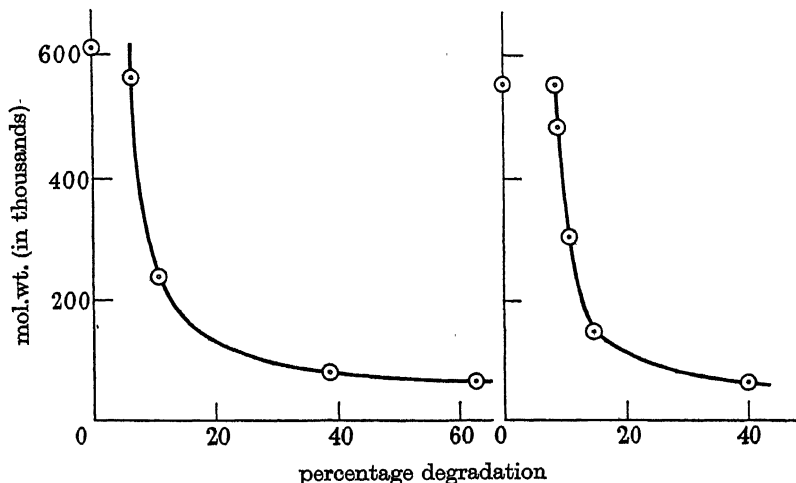


FIGURE 8

FIGURE 9

FIGURE 8. Effect of degradation on the molecular weight of 1/410 copolymer.

FIGURE 9. Effect of degradation on the molecular weight of 1/40 copolymer.

TABLE 4. MOLECULAR WEIGHTS ON DEGRADATION OF 1/410 COPOLYMER

temperature = 220° C		
time of degradation	% degradation	mol.wt.
0	0	617,000
50	3	560,000
120	7	234,000
280	35	80,000
580	59	64,000

TABLE 5. MOLECULAR WEIGHTS ON DEGRADATION OF 1/40 COPOLYMER

temperature = 220° C		
time of degradation	% degradation	mol.wt.
0	0	547,000
90	1	475,000
190	2.7	247,000
290	8	140,000
590	33	66,000

These results again indicate that breaks occur at the acrylonitrile units, leaving residues of pure methyl methacrylate chains which of course must degrade by reverse polymerization and show no fall in molecular weight.

So far it has been established that the role of the acrylonitrile is to create weak links which break easily.

In measuring the amounts of monomer produced, initial volatiles were not allowed for. Figures 8 and 9 show that they amount to about 5 to 10 % of the original weight of polymer since the curves tend to this value and not to zero degradation.

An approximate theory

The rate-time curves, figures 6 and 7, show that small amounts of monomer are always produced in the reaction. At the beginning of the reaction it is probable that most of the monomer comes from degradation of the ends of the chains to the first acrylonitrile unit, where it is stopped. The rate of production of monomer then increases with time until it reaches a maximum.

The initial breaks will almost certainly not occur within the acrylonitrile units since they do not do so in the case of pure polyacrylonitrile which does not degrade. The break cannot take place at a methyl methacrylate-methyl methacrylate link since this does not occur in pure polymethyl methacrylate. It must therefore be assumed that the break occurs at the methyl methacrylate-acrylonitrile link, and where this break occurs two free radicals must be formed. One of these will be terminated by an acrylonitrile unit and the other by a methyl methacrylate unit. The acrylonitrile is not degradable otherwise the copolymer would degrade completely to monomer in the same way as pure polymethyl methacrylate. This acrylonitrile must therefore become stable in some way while the methyl methacrylate end degrades to the next acrylonitrile unit, which on the average is 410 units away. Therefore for each break which occurs in a chain, 410 units of methyl methacrylate are produced in the 1/410 copolymer and 40 are produced in the 1/40 copolymer sample.

Since ten times as much monomer is produced per break from the 1/410 sample, the molecular weight against percentage degradation to monomer curve would be expected to fall much more steeply in the case of the 1/40 sample. Reference to figures 8 and 9 shows that this is indeed the case.

Assuming that (i) breaks occur at the acrylonitrile units, and (ii) x methyl methacrylate units are produced at each break, the number of breaks which occur can be worked out in the following way from the knowledge of the fall in molecular weight with time. If $(D.P.)_0$ = original degree of polymerization, and n = number of breaks per molecule, i.e. $n + 1$ = number of fragments produced from each molecule, x = the ratio, methyl methacrylate/acrylonitrile in the original molecules,

$D.P.$ = degraded degree of polymerization,

then
$$D.P. = \frac{\text{no. of units present in involatile fragments}}{\text{no. of fragments produced from each molecule}}$$

i.e.
$$D.P. = \frac{(D.P.)_0 - xn}{n + 1}.$$

This formula is only true in the initial stages of the reaction since two fragments have been assumed to have been formed as a result of each break. If the acrylonitrile

group nearest the end of the molecule breaks, however, and if the methyl methacrylate units eliminated are those between this acrylonitrile and the end of the molecule then only one involatile residue will be left from such a break. With the 1/410 sample there are on the average 15 acrylonitrile units in each chain. The probability therefore of such a bond breaking is initially 1 in 15. When, however, one break has occurred on the average in each molecule the probability of only one involatile being left will be about 1 in 7. As degradation proceeds the probability will become greater and the formula will become less and less applicable. It should, however, give a good approximation at least in the initial stages of the reaction, of the number of breaks which have occurred on the average in each molecule.

For the 1/400 polymer the formula becomes,

$$D.P. = \frac{6170 - 410n}{n + 1},$$

so that, knowing the molecular weight of the degraded sample, the average number of bonds broken may be calculated.

The results obtained for the 1/410 and 1/40 copolymer samples are calculated from tables 4 and 5 and shown in figures 10 and 11.

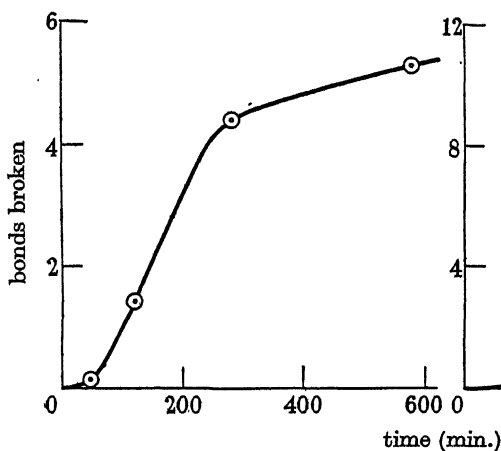


FIGURE 10

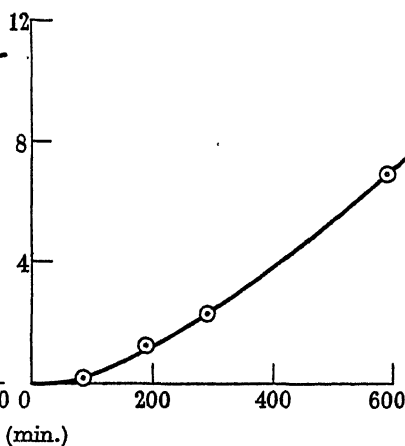


FIGURE 11

FIGURE 10. Rate of breaking of bonds in 1/410 copolymer at 220° C.

FIGURE 11. Rate of breaking of bonds in 1/40 copolymer at 220° C.

There is apparently a very close agreement between the curves in figures 10 and 11 and those in figures 6 and 7 although the two sets of results were obtained in totally different ways. In figure 10 the apparent induction period and the increase and subsequent decrease in rate can be seen and these agree almost exactly in time with the similar evidence in figure 6. Figure 11 shows that even after 590 min. at 220° C, the 1/40 copolymer has not reached the maximum rate of degradation as indicated by a point of inflexion on the bonds broken-time curve. Again this agrees with the rate measurements shown in figure 7 where after 590 min. the rate is just about the maximum.

590 min. at 220° C corresponds to 39 % degradation of this polymer sample, and since even after that time the maximum rate of breaking of bonds has not been reached, additional evidence is thereby obtained that the molecular weight, as shown in figure 9, would fall to a much lower value at higher degrees of degradation.

Good verification of the theory has been obtained as follows.

From the bonds broken-time curve, the rate of production of monomer may be obtained at any time from the slope, and it is easiest to draw a tangent to the curve at the point of greatest slope which according to the theory, gives the maximum rate of production of monomer.

From the bonds broken-time curve, the maximum rate of breaking of bonds is

$$\begin{aligned} 1 \text{ bond/molecule/43 min.} &= 410 \text{ methyl methacrylate units/molecule/43 min.} \\ &= 9.53 \text{ methyl methacrylate units/molecule/min.} \end{aligned}$$

Therefore the average fraction of each molecule which is lost per minute as monomer

$$= 9.53/6170.$$

0.3 g. of polymethyl methacrylate were originally present, therefore $\frac{9.53 \times 0.3}{6170}$ g. of monomer should be produced per minute,

$$\begin{aligned} &= \frac{9.53 \times 0.3 \times 60}{6170} \text{ g./hr.} \\ &= \frac{9.53 \times 0.3 \times 60}{6170 \times 0.94} \text{ ml./hr.} \\ &= 0.0296 \text{ ml./hr.} \end{aligned}$$

The maximum rate, from figure 6, is given by $V = 0.475$, and $V_0 = 0.365$, which gives $(V^2 - V_0^2)/V_0^2 = 0.71$ and according to figure 4, part I, this is equivalent to a rate of production of monomer of 0.0305 ml./hr.

Such good agreement is not obtained with all points on the curves, but considering the factors involved, it seems satisfactory even to have obtained agreement to within a factor of two or three.

The induction period is rather difficult to understand, but an explanation appears to be that the acrylonitrile units (where the breaks occur) tend to be protected from attack within the coils of the larger molecules, so that the rate will only increase when the molecules have been somewhat broken down. Another possibility is that the two fragments formed at a break do not become free radicals but disproportionate to molecules, at least one of which has the potentiality to degrade at some future time. The rate will only increase when a high concentration of these molecules has been built up.

Energy of activation

It has been established that the degradation of copolymers proceeds by breaking at the acrylonitrile units, followed by the degradation of part of the chain. This reaction is very slow in the initial stages but becomes faster as the chains are partially broken down. Although breaking of acrylonitrile bonds is slow, in the initial stages a certain amount of monomer is produced by degradation from the ends of the

chains, as in the case of pure methyl methacrylate, but only to the first acrylonitrile unit. Once all the chains have degraded to the first acrylonitrile unit, each chain will have at one end, either a benzoyl or a phenyl group, and at the other, an acrylonitrile unit. Henceforth, all degradation must take place by chain scission at the acrylonitrile bonds.

It would be of interest to discover whether degradation by the breaking of these acrylonitrile bonds requires more or less energy than the normal degradation through the ends of the molecules.

The energies of activation at various stages of the reaction, which are shown in table 6, were measured by the usual methods.

TABLE 6. ENERGY OF ACTIVATION OF DEGRADATION OF COPOLYMERS

1/410 copolymer		1/40 copolymer	
% degradation	<i>E</i>	% degradation	<i>E</i>
2	32,000	7	29,200
10	30,700	16	25,000
31	19,700	31	19,200

The initial value of about 30,000 is probably due to the degradation from the ends of the chains. As the copolymer breaks up into smaller fragments the energy required to break the acrylonitrile bonds becomes less, probably because the bonds are more vulnerable to attack in the smaller molecules.

Thanks are due to I.C.I. Ltd. (Dyestuffs Division), for samples of 1,4 diamino anthraquinone, to I.C.I. Ltd. (Plastics Division), for certain polymer samples, and to the Carnegie Trust for financial assistance to one of us (N.G.).

The experimental work was done in the Chemistry Department of the University of Aberdeen during the tenure by one of us (N.G.) of a Carnegie Scholarship.

REFERENCES

- Alfrey, T., Merz, E. & Mark, H. 1947 *J. Polymer. Sci.* **1**, 37.
 Alfrey, T. & Price, C. C. 1947 *J. Polymer. Sci.* **2**, 101.
 Lewis, F. M., Mayo, F. R. & Hulse, W. F. 1945 *J. Amer. Chem. Soc.* **67**, 1701.
 Norrish, R. G. W. & Brookman, E. F. 1939 *Proc. Roy. Soc. A*, **7**, 147.
 Price, C. C. 1946 *J. Polymer. Sci.* **1**, 83.

The angular distribution of γ -radiation from light nuclei

I. Experimental

BY S. DEVONS, PH.D. AND M. G. N. HINE,
Cavendish Laboratory, University of Cambridge

(Communicated by O. R. Frisch, F.R.S.—Received 28 January 1949—

Revised 19 April 1949—Read 16 June 1949)

The experimental procedure used in studying the angular distribution of γ -radiation produced by the proton bombardment of some light nuclei is described. Results are given for the radiation following proton capture by the nuclei Li^7 , Be^8 , C^{12} , C^{13} and for the radiation from the residual O^{18*} nucleus produced in the reaction $\text{F}^{19}(p, \alpha) \text{O}^{18*}$. Proton energies between 300 and 1000 keV were available. The distribution of the radiation from Li^7 was investigated over the whole range of energies; in the other cases the distribution was measured at energies corresponding to resonance maxima for the processes.

The main results are:

reaction	proton energy (keV)	angular distribution
$\text{Li}^7(p, \gamma) \text{Be}^8$	440	$1 + 0.05 \cos \theta$
$\text{Be}^8(p, \gamma) \text{B}^{10}$	960	$1 + 0.09 \sin^2 \theta$
$\text{C}^{12}(p, \gamma) \text{N}^{13}$	450	isotropic
$\text{C}^{13}(p, \gamma) \text{N}^{14}$	550	isotropic
$\text{F}^{19}(p, \alpha) \text{O}^{18*}$	330, 470 and 670	isotropic
	590	$1 + 0.2 \cos^2 \theta$
	870	$1 + 0.1 \cos^2 \theta$

INTRODUCTION

In the study of γ -transitions between the energy levels of medium and heavy nuclei, information about the multipole nature of the γ -radiation can be obtained from measurements of the internal conversion coefficients for the emission of K , L , etc., orbital electrons, and this information, supplemented by measurement of the transition energy, and sometimes also of the half-life, is of value in establishing the properties of the nuclear states between which a transition occurs. Such procedures are usually inapplicable to the study of γ -transitions in light nuclei, and particularly those occurring in competition with nucleon emission, for the following reasons. The internal conversion coefficients are normally much smaller in light than in heavy nuclei. The commonly occurring transition energies are also somewhat larger for light nuclei, which is one of the features responsible for small conversion coefficients. The radiating nucleus may be recoiling as a result of the nuclear transmutation process in which it was formed in an excited state, so that it cannot be reliably assumed that the nucleus is surrounded by the normal electron configuration. In addition, in the particular case of γ -radiation associated with nuclear reactions and emitted during bombardment, the experimental conditions are not particularly suitable for precise measurement of internal conversion coefficients.

Information about the multipole nature of a γ -transition can, in principle, be obtained from a study of the electron-positron pairs produced by internal conversion

in the field of the radiating nucleus. The probability of this process is almost independent of the nuclear charge, and of the presence or absence of the extranuclear electrons, and increases with increasing transition energy (Jaeger & Hulme 1935). Such conversion coefficients are, however, quite small even in favourable cases, and their dependence on multipolarity is not so pronounced as to make the method easy to employ. The spatial distribution of the electron pairs is also dependent on the multipolarity of the transition, but again the dependence is not very marked (Horton 1948).

In view of the difficulties in applying the above-mentioned methods to the study of γ -radiation from light nuclei, we have made some investigations of the extent to which the properties of a particular radiative transition can be inferred from measurements of the angular distribution (with respect to the incident beam of particles) of the γ -radiation emitted by excited nuclei in some reactions involving light nuclei. The present paper describes the experimental procedure and results. The theoretical interpretation of the results is dealt with in the following paper, part II.

The interpretation of such measurements is least ambiguous in the case of γ -radiation emitted directly after particle capture, and most of the investigations have been of processes of this type (proton capture by Li^7 , Be^9 , C^{12} , C^{13} and F^{19}). When the γ -radiation is emitted from a residual nucleus, produced in an excited state as a result of a nuclear transmutation or by inelastic scattering, the interpretation of any observed angular correlation will usually require considerable additional information about the particular nuclear processes involved. One such case has been studied, the reaction $\text{F}^{19}(p, \alpha)\text{O}^{16*}$.

A few measurements of the angular distribution of γ -radiation in some of these reactions have already been reported, but the results obtained do not seem to have been sufficiently precise or extensive for substantial deductions to have been made from them. (Ageno, Amaldi, Bocciairelli & Trabacchi 1941; Van Allen & Smith 1941; Jacobs, Malmberg & Wahl 1948; Fowler, Lauritsen & Lauritsen 1948).

EXPERIMENTAL TECHNIQUE

In many instances the expected and observed degree of anisotropy in the radiation pattern is quite small, so that one of the principal requirements of the experimental arrangement is that all extraneous sources of anisotropy should be eliminated. In practice it is usually necessary, ultimately, to compromise between this requirement and that of adequate intensity, and also, but less important, between the latter and the angular resolution used.

The arrangement used is shown in figure 1. A parallel beam of protons produced by a Cockcroft-Walton type accelerator was deflected through 90° and brought to a focus by the magnet H . The adjustable horizontal slit S_1 was used for energy resolution, the highest resolution employed being about 0.3 %, corresponding to a vertical slit width of 0.4 mm. The circular apertures S_2 , S_3 define the beam incident on the target T which is a deposit, obtained by evaporation *in vacuo*, of the material under investigation on a thin brass foil (0.2 mm. thick). This foil is mounted on a water-cooled, thin-walled copper tube arranged so that the front side of the foil passes accurately through the axis about which the frame supporting the counters rotates.

to such sources seldom exceeded a few percent of the genuine effect and was usually much less.

The radius of the aperture S_3 and the distance of the movable counters from the axis of rotation govern the geometrical imperfection of the arrangement, and these were determined in particular instances by the need for adequate intensity. The size of the beam spot on the target ranged, in different experiments, from about 1 to 3 mm. radius, and the distance of the inner counters from the axis from 7.5 to 15 cm. The angular resolution varied therefore from about $\pm 3^\circ$ to $\pm 6^\circ$ at 90° to the incident beam, and from $\pm 8^\circ$ to $\pm 16^\circ$ at 180° or 0° to the beam.

The advantage gained by using two pairs of counters, apart from the obvious increase in statistical accuracy, lies in the fact that accidental anisotropy introduced by small displacements of the beam, in the plane of figure 1, are eliminated to the first order if the average distribution as registered by both pairs of counters is used in estimating the real angular distribution. The target support could be rotated in the course of a series of measurements, and its position relative to that of the counters could be arranged so that absorption of the γ -radiation by the copper tube supporting the target backing was unimportant. The absorption in the target backing itself never amounted to more than 1 % and was allowed for where significant.

Angular distributions were measured by observing the ratio of the number of coincidences, in each counter pair, cC , to the counts in the monitor M as a function of the position of the former. Check experiments showed that the intensity recorded by the monitor was predominantly proportional to the radiation from the target, although a small correction, always less than 3 %, had to be made for other sources of radiation, in particular that arising from proton bombardment of the apertures S_1 , S_2 . The mean intensity of the radiation was usually held constant to within $\pm 10\%$ during measurement of the angular distribution at a particular proton energy.

A resolving time of $1.5 \mu\text{sec.}$ was used in the coincidence circuits, and the random 'coincidences' were estimated by feeding the pulses from one pair of counters into a supplementary coincidence circuit with a much larger resolving time ($20 \mu\text{sec.}$).† The procedure was necessary since the random 'coincidence' rate could not be computed from the mean counting rate in each counter of a coincidence pair, on account of the inherent modulation (mainly at 200 c./sec.) of the beam current on the target caused by the a.c. ripple of the high-voltage equipment and other sources of current fluctuation. The number of genuine coincidence N is then given by the expression

$$N = N_S - \{(N_L - N_S)\tau_S/(\tau_L - \tau_S)\},$$

where N_S and N_L are the total numbers recorded by the channels with short and long resolving times and τ_S , τ_L are the respective resolving times themselves. The correction due to random coincidences was usually of the order of 1 % of genuine count, or less. The errors indicated in the individual results given below are derived simply from the normal statistical fluctuations in the numbers of counts, both in the coincidence counters and monitor. The compromise between ideal geometry and

† We are indebted to Mr H. G. Hereward for drawing our attention to this procedure.

maximum intensity was usually made so that other sources of error were substantially smaller than the random statistical one, but in some cases (where the radiation intensity is particularly weak) the two sources of error are comparable.

CORRECTIONS FOR CENTRE OF GRAVITY MOTION

For a comparison of the experimental results with theoretical predictions it is necessary to deduce from the measurements the angular distribution in a frame of reference at rest with respect to the centre of gravity of the radiating nucleus ('centre of gravity frame'). If $f(\theta)$ is the relative intensity of radiation (i.e. number of quanta) per unit solid angle in the direction θ to the proton beam, as measured in the laboratory frame of reference, and v is the velocity of recoil of the radiating nucleus, then a Lorentz transformation gives for the intensity per unit solid angle, $g(\theta')$, in the centre of gravity frame of reference:

$$g(\theta') = f(\theta) \frac{(1 - v^2/c^2)}{(1 + v/c \cdot \cos \theta')^2},$$

where θ' is the transformed value of θ given by

$$\cos \theta = \frac{\cos \theta' + v/c}{1 + v/c \cdot \cos \theta'}.$$

Since v/c is of the order 0.01 or less in all the cases investigated, terms of higher power than the first in v/c are neglected. With this approximation we obtain for the distribution in the centre of gravity frame, the expression

$$g(\theta) = \{f(\theta - v/c \cdot \sin \theta)\} \{1 - 2v/c \cdot \cos \theta\}.$$

A further correction that has to be applied in practice arises from the Doppler frequency shift associated with the recoil motion and the effect of this frequency shift on the efficiency of detection of γ -radiation. If E is the quantum energy (independent of angle) in the centre of gravity frame, then in the laboratory frame the energy is $E(1 + v/c \cos \theta)$. For such small relative changes in quantum energy we have assumed that the variation in efficiency of γ -ray detection can be estimated as follows. In a curve showing the number of coincidences between a pair of counters (such as c_1, C_1) as a function of the thickness of absorber between them (and with a constant thick radiator) it appears from experimental observations that: (a) the number of coincidences per quantum at zero absorber thickness is linearly proportional to the quantum energy, and (b) the *shape* of the absorption curve, allowing for the change in scale of absorber thickness with energy, is roughly independent of the quantum energy (Fowler *et al.* 1948). If the number of coincidences as a function of absorber thickness is represented by $N(t)$, then with the above assumptions, the change, δN , in N due to a small change δE in the quantum energy is given by

$$\delta N/N = (\delta E/E) (1 - t/N \cdot dN/dt \cdot \delta E/E).$$

The appropriate values of dN/dt are obtained from measurements made with the particular geometry used in the main experiments for one or two quantum energies

and comparison of these results with similar measurements made by other investigators (Bleuler & Zunti 1946; Fowler *et al.* 1948). The maximum correction that was necessary for Doppler frequency shift amounted to $\sim 1.5\%$.

EXPERIMENTAL RESULTS

All the measurements reported below have been made with protons as the bombarding particles and with target materials containing the naturally occurring isotopic abundance ratios.

Lithium† ($\text{Li}^7(p, \gamma)\text{Be}^8$)

The γ -radiation studied arises from the reaction $\text{Li}^7(p, \gamma)$. The cross-section for this reaction shows a pronounced resonance at a proton energy of 440 keV, but has an appreciable and approximately constant value for proton energies up to 1.0 MeV which cannot be attributed to the same level of the compound nucleus as that responsible for the resonance at 440 keV. Making the assumption that for proton energies in the neighbourhood of 440 keV the reaction can be treated as arising from a single level of compound nucleus, several investigators have studied the variation and absolute magnitude of the cross-section‡ and deduced from their results the following properties of the level (Bonner & Evans 1948; Fowler *et al.* 1948): proton width (Γ_p) $\approx$ total width (Γ) = 12 keV. γ -radiation width (Γ_γ) = $9g^{-1}$ eV; γ -energy (at 440 keV) 17.5, 14.7 MeV (g is the undetermined statistical factor

$$(2J + 1)/(2j_A + 1)(2j_B + 1)).$$

Recent measurements of the γ -radiation spectrum (Walker & McDaniel 1948) indicate that the ratio of the intensities of the 14.7 and 17.5 MeV components is not independent of proton energy.

The first measurements of the angular distribution of the radiation were made using a thick target (fused LiOH) and a proton energy of 500 keV. Under these circumstances the radiation is principally from the 440 keV resonance, although as will be seen later even an approximate distinction between the resonance and background radiation is apt to be misleading. The average of the results obtained from several sets of observations is shown in figure 2. The solid curve *a* shows the best fit to the experimental points (corrected only for absorption by the target support) of the type $1 + \alpha \cos \theta$. The dotted curves *b* and *c* illustrate the magnitude of the corrections that have to be applied to allow for recoil motion. Curve *c* is the one to be compared with theory. Figure 3 shows the typical results obtained from two individual sets of observations, one with an aluminium absorber of 8 mm. thickness, the other 2 mm. In order to check more thoroughly the effect of the Doppler frequency shift on detection efficiency the absorber between the counters was increased to 22.5 mm., and the relative intensities of the radiation in the forward (15°) and backward (145°) directions were remeasured. With this absorber thickness the effect of the 14.7 MeV component is almost entirely eliminated, and the detection

† Some preliminary results of these investigations have been published elsewhere (Devons & Hine 1948).

‡ The angle at which the radiation is measured is not usually specified.

efficiency for the 17.5 MeV component more sensitive to exact γ -ray energy than with the thinner absorbers. The ratio $I(15^\circ)/I(145^\circ)$ for the 8 mm. absorber was found to be 1.11 ± 0.005 , and this ratio increased to 1.15 ± 0.01 for the 22.5 mm. absorber. Correction for Doppler frequency shift reduces the first ratio to 1.10 and the second to 1.09. Thus the increased asymmetry observed with the thicker absorber can be attributed entirely to the Doppler frequency shift associated with a Be^8 nucleus moving with the full recoil velocity. This result is consistent with the short half-life to be expected for the compound Be^8 nucleus (of the order 10^{-19} sec., according to the measured width of the resonance level). It is possible, of course, that part of the observed change in the angular distribution is due to the elimination of the effect of the 14.7 MeV component, but in view of the difficulty of checking this point we have made the simple and plausible assumption that the angular distribution of the two components is similar.

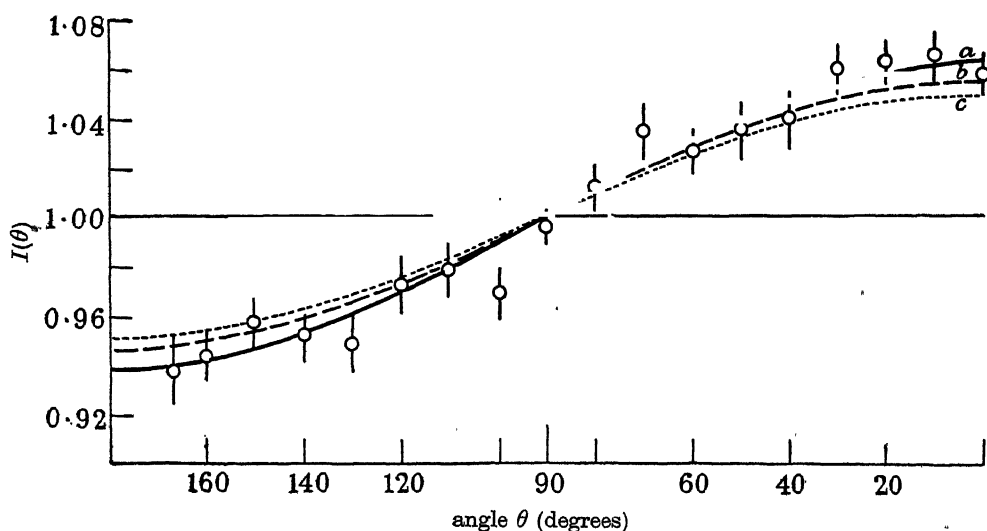


FIGURE 2. Angular distribution of γ -radiation from a thick lithium target. Proton energy: 500 keV. Absorber thickness: 8 mm.

The observed asymmetry indicates the presence of interference between two states (of the compound nucleus) of opposite parity (see part II), and since one of these states corresponds to the sharp resonance at 440 keV, the asymmetry should be a pronounced function of proton energy. This variation was examined, first, by measuring the ratio of the intensities of the 'forward' (15°) and 'backward' (145°) radiation for proton energies of 250 to 950 keV, using thin lithium targets of thicknesses corresponding to stopping powers of between 6 and 20 keV, and also by more detailed measurement of the angular distribution at 650 and 900 keV. The results of the first set of measurements are shown in figures 4 and 5, the latter showing the detailed measurements in the neighbourhood of the resonance maximum. In all cases the proton energy plotted is the effective mean energy averaged through the target thickness. This is not exactly equal to the incident proton energy less half

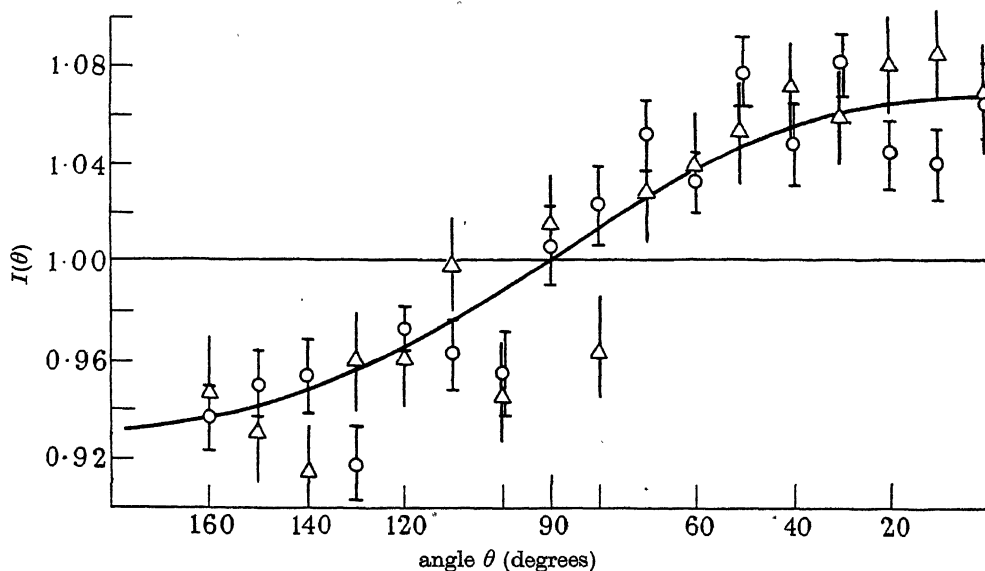


FIGURE 3. Angular distribution of the $\text{Li}^7(p, \gamma)$ radiation with thick and thin absorbers.
Absorbers: $\circ = 0.2$ cm. Al; $\triangle = 0.8$ cm. Al.

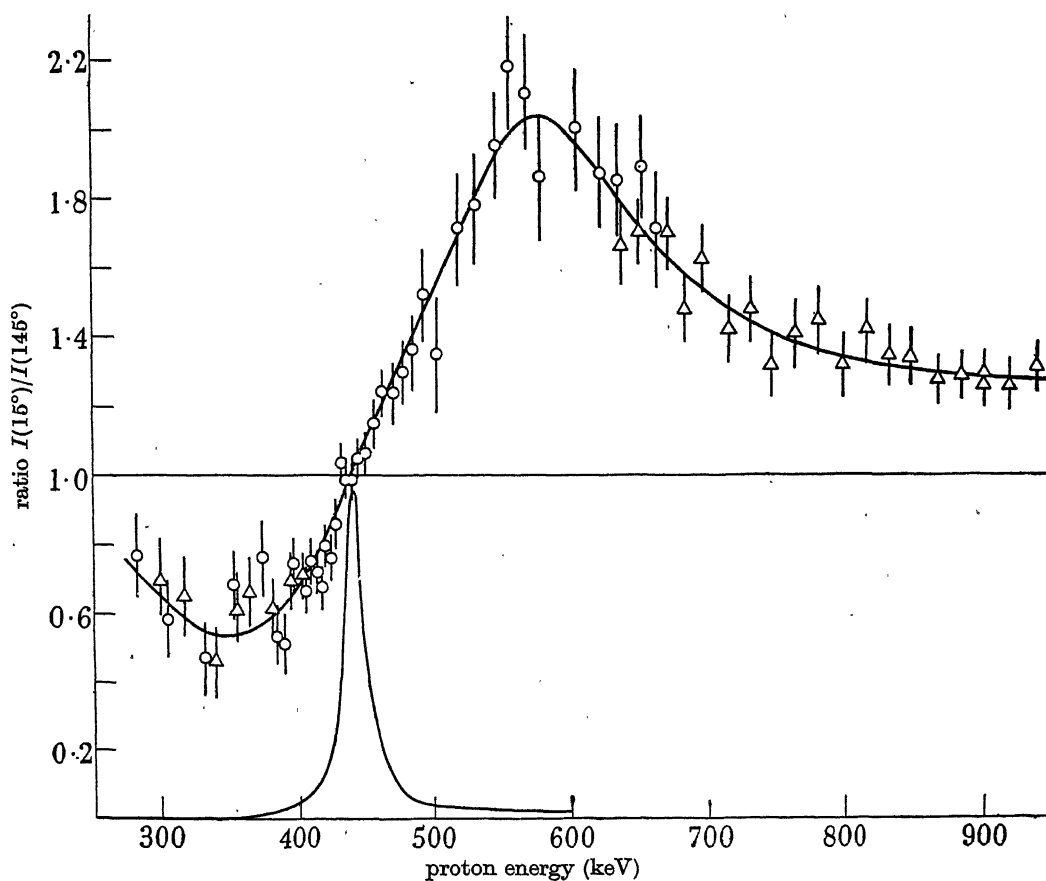


FIGURE 4. Variation of asymmetry with proton energy.
Target thickness $\circ = 12$ keV at 500 keV; $\triangle = 20$ keV at 500 keV.

the energy loss in traversing the target, since the absolute intensity of radiation, at any angle, varies with energy, and particularly rapidly in the neighbourhood of 440 keV, so that the effective mean energy is not equal to the mean energy of the protons traversing the target. In estimating the correction to be applied to the incident proton energy, the value 12 keV was adopted for the total width of the resonance level.† With the thinnest targets used the simple correction of half the

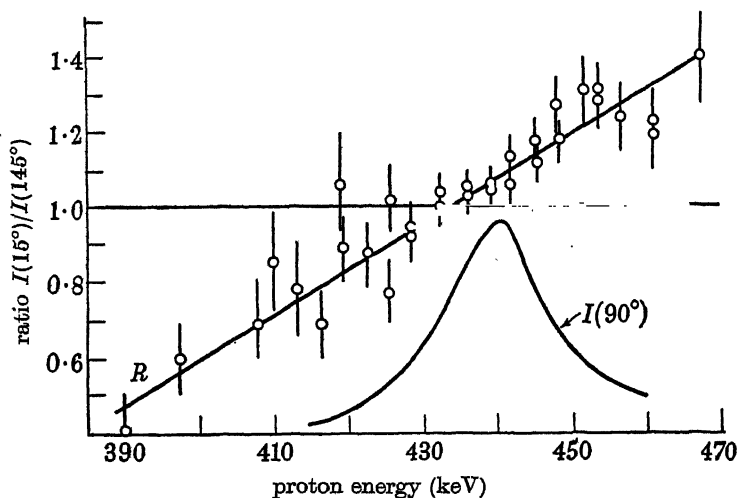


FIGURE 5. Variation of asymmetry near 440 keV. Target thickness ~ 6 keV.

energy loss in the target does not differ substantially from the more precise correction. The angular distribution for proton energies of 650 and 890 keV are shown in figures 6 *a* and *b*. The solid lines represent functions of the type $1 + \alpha \cos \theta + \beta \cos^2 \theta$ fitted to the experimental points after correction for recoil motion and Doppler shift have been made. At the time that these measurements were made there was some indication that the energy of the γ -radiation was substantially different for the 440 keV resonance and for higher energies (Fowler *et al.* 1948), a fact that would conflict with the interpretation of the observed asymmetry as due to interference between two levels of the compound nucleus from which γ -transitions of the same energy could occur. A qualitative check was therefore made for the presence of an appreciable component of 17.5 MeV energy with proton energies some distance from the resonance. This was done by measuring the excitation function for the radiation with two pairs of counters, one having an absorber thickness of 8 mm. and the other 22.5 mm. The first pair detects the 14.7 and 17.5 MeV components with comparable efficiency and the second predominantly the latter. From the results (figure 7) the ratio of the two components can be estimated to be approximately equal away from the resonance, and the 17 MeV component to be about four times as intense as the 14 MeV component at the resonance. This result is in accord with more quantitative spectrometer measurements of the radiation that have been made recently (Walker & McDaniel 1948).

† This refers to the energy width as measured in the laboratory frame of reference with proton bombardment. The real width is seven-eighths of this.

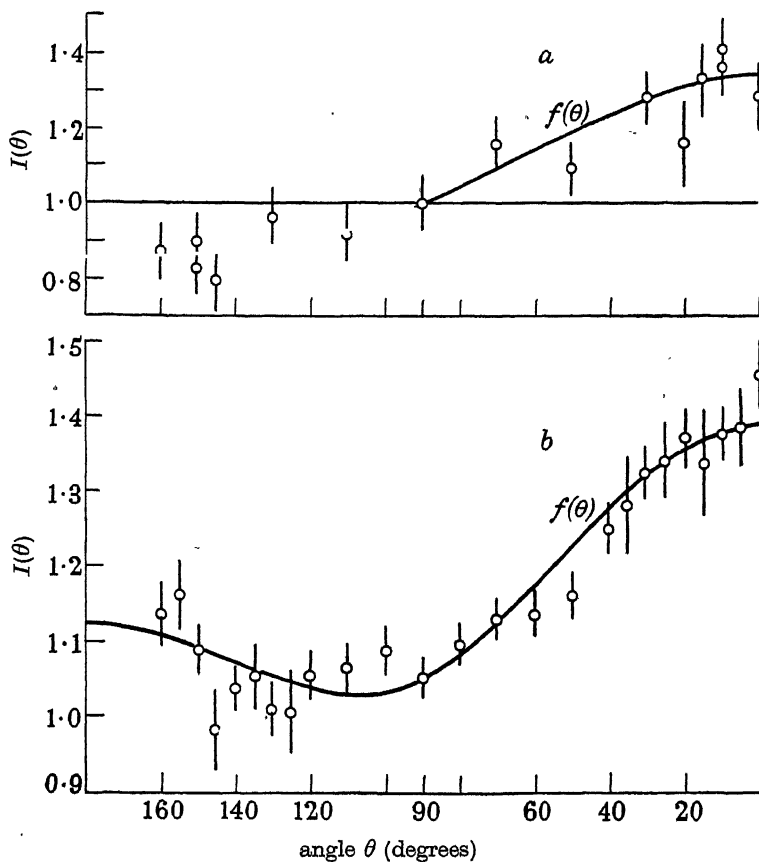


FIGURE 6. Angular distribution of $\text{Li}^7(p, \gamma)$ radiation for proton energies of
 a 650 keV, $f(\theta) = 1 + 0.25 \cos \theta + 0.1 \cos^2 \theta$.
 b 890 keV, $f(\theta) = 1 + 0.12 \cos \theta + 0.20 \cos^2 \theta$.

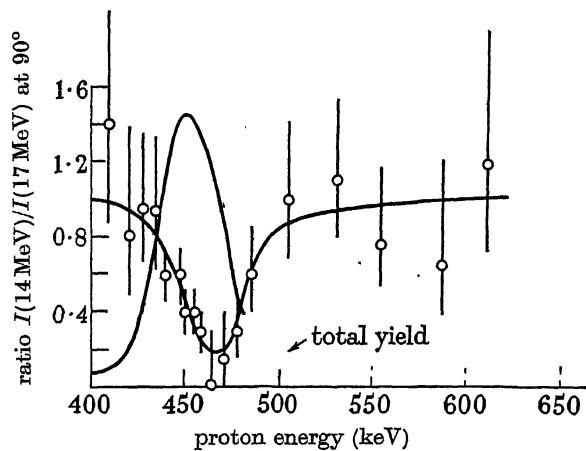


FIGURE 7. Variation of spectrum of the $\text{Li}^7(p, \gamma)$ radiation with proton energy.

A quantity which is significant in providing a quantitative theoretical interpretation of these results is the proton energy E_0 for which the radiation is exactly isotropic. From figure 5 this energy can be seen to be somewhat different from that of the resonance maximum, E_r , the difference $(E_r - E_0)$ being about 6 keV. An alternative estimate of this quantity can be obtained from the measured anisotropy of the radiation from a thick target (figure 2). Assuming that the ratio $I(15^\circ)/I(145^\circ)$ varies linearly with proton energy in the region 400 to 500 keV (cf. figure 5), the asymmetry for a thick target and bombarding energy of 500 keV is due almost entirely to difference between E_r and E_0 , since the shape of the excitation function is almost symmetrical about the energy E_r in this range of energies. Using the observed variation of the ratio $I(15^\circ)/I(145^\circ)$ shown in figure 5 and the excitation curve given by Fowler *et al.* (1948), the asymmetry shown in figure 2 is found to correspond to a value of $E_r - E_0$ of 7 ± 2 keV, which is consistent with the approximate value obtained directly.

Beryllium ($\text{Be}^9(p, \gamma)\text{B}^{10}$)

Resonances for the capture of protons by Be are well established at proton energies of 990 and 1080 keV with the following properties (Fowler *et al.* 1948):

$E_r = 990$ keV: $\Gamma = 94$ keV, $\Gamma_\gamma \approx 12.5g^{-1}\text{eV}$, γ -energy 7.4 MeV.

$E_r = 1080$: $\Gamma = 4$ keV, $\Gamma_\gamma \approx 0.8g^{-1}\text{eV}$, γ -energy 6.7, 0.8 MeV.

There is possibly an additional weakly excited level corresponding to a proton energy of 880 keV.

We have measured the angular distribution of the radiation from the 990 keV level, using as targets thin evaporated films of beryllium whose stopping power for protons was about 10 keV (at 1 MeV). The aluminium absorbers between the counters was 0.5 cm. thick. The average of several sets of observations is shown in figure 8. The departure from isotropy is very small and although it appears to be significant it should be remarked that the magnitude of the anisotropy varied appreciably amongst the individual sets of observations, indicating sources of error other than the normal statistical fluctuations. These arise, no doubt, from the somewhat degraded geometrical arrangements (larger beam defining apertures) adopted in order to obtain maximum intensity. The results are consistent with the interpretation of the process as arising predominantly from a single level of the

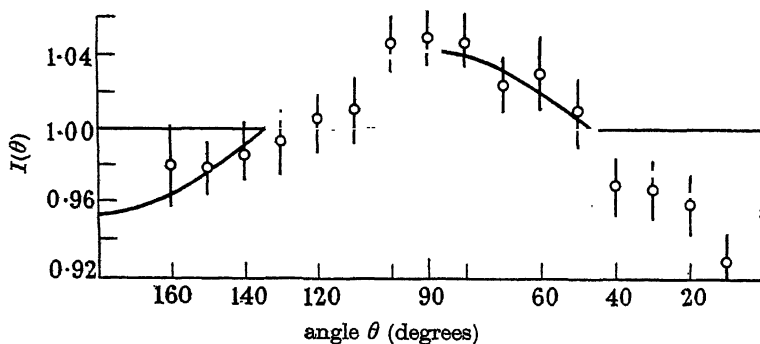


FIGURE 8. Angular distribution of $\text{Be}^9(p, \gamma)$ radiation.

compound B^{10} nucleus. Measurable interference between the levels at 990 and 1080 keV would not be expected if the energies of the γ -transitions from the two states are different, and in any case the width of the higher level is so small that its contribution at the proton energy employed would be barely measurable.

In the course of these measurements we had occasion to investigate the excitation curve for the $Be^9(p, \gamma)$ reaction, both to establish the energy scale and to determine the target thickness. With every Be target studied in the region 800 to 900 keV we observed some evidence of the reported weak resonance at the proton energy of 880 keV, and we consider it unlikely that this was due to fluorine contamination. Some measurements of the angular distribution were made with protons of this energy, but the statistical accuracy was not sufficient to obtain a more precise conclusion than that the departure from isotropy, if any, was less than 10 %.

Carbon

Two resonances for the capture of protons by carbon are well established with the following properties (Fowler *et al.* 1948):

	E_r (keV)	Γ (keV)	Γ_γ (eV)	energy of γ -radiation (MeV)
C^{12}	450	35	$0.63g^{-1}$	2.3
C^{13}	550	40	$15g^{-1}$	2.3, 5.8 and 8.1

The angular distribution of the radiation was measured using a 'thick' carbon target (a deposit of soot).

(i) $C^{12}(p, \gamma)N^{13}$

A proton energy of 500 keV was used. At this energy the radiation from the tail of the 550 keV resonance in C^{13} makes a small contribution to the measured effect, but since this latter radiation is practically isotropic (*vide infra*), this circumstance does not seriously affect the conclusions that can be drawn from the observations.

An absorber of only 1 mm. thickness was used between the counters in order to obtain reasonable counting rates, and with such a thin absorber there was a small contribution to the counting rate from the annihilation radiation produced by

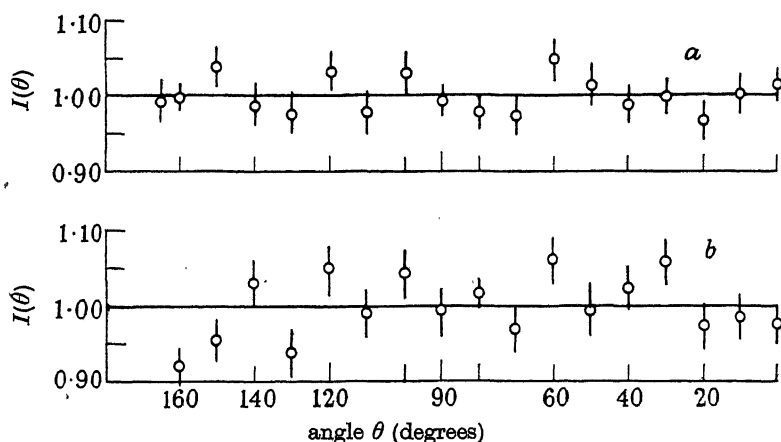


FIGURE 9. Angular distribution of radiation from carbon. *a* $C^{12}(p, \gamma)N^{13}$; *b* $C^{13}(p, \gamma)N^{14}$.

positrons from the N^{13} formed in the reaction. A small correction was made for this contribution, determined by observing the counting rate immediately after the proton beam had been switched off. It was also necessary to use additional lead shielding around the monitoring counter to ensure that the contribution to its counting rate from the annihilation radiation was negligible, in comparison with the more energetic capture radiation. The results, corrected for absorption in the target support and Doppler shift are shown in figure 9*a*. There is no significant departure from isotropy. A distribution isotropic in the centre of gravity frame would become in the laboratory frame of reference, $1 + 0.005 \cos \theta$.

(ii) $C^{13}(p, \gamma)N^{14}$

Protons of 625 keV energy and aluminium absorbers between the counters 4 mm. thick were used. Under these conditions the contribution from the C^{12} resonance was negligible (as confirmed experimentally) and the sensitivity of the arrangement to the 2.3 MeV component extremely small. Effectively, then, the measurements determine the angular distribution of the combined 8.1 and 5.8 MeV components, of which the first is certainly, and the second probably, a direct transition from the state of N^{14} responsible for the capture resonance. The results (figure 9*b*), with the usual corrections applied, indicate no significant departure from isotropy.

Fluorine ($F^{19}(p, \alpha)O^{16*}, O^{16*} \rightarrow O^{16} + \gamma$)

The variation of the cross-section of this reaction with proton energy has been extensively studied and there are well-defined resonance maxima, for proton energies of less than 1 MeV, at 338, 470, 590, 665, 860 and 920 keV, and several much more weakly excited levels (Bonner & Evans 1948; Bennett, Bonner, Mandeville & Watt 1946). Recent measurements of the γ -spectrum from the residual O^{16} nucleus, and of the energy of the short-range α -particles, indicate that this nucleus may be produced in one of two excited states which can emit γ -radiation, the excitation energies of these states being 6.1 and 7.0 MeV. For proton energies less than 700 keV the radiation is almost entirely from the lower of these two states. The angular distribution of the radiation has been measured for the first four of the resonance levels mentioned above, using in each case thin targets made by evaporating CaF_2 *in vacuo*, and aluminium absorbers between the counters 4 mm. thick.

No correction for Doppler frequency shift or centre of gravity motion has been applied to any of the results in the case of fluorine γ -radiation. In addition to the primary recoil of the compound Ne^{20} nucleus one should also consider the recoil of the excited O^{16} nucleus left after the emission of an α -particle. For an α -particle of 2 MeV energy the recoil velocity of the O^{16} nucleus is about 4×10^8 cm./sec. If the half-life for the subsequent γ -transition is of the order 10^{-13} sec. or more (i.e. radiation width of 10^{-2} eV or less) it is probable that most of the recoil momentum of the O^{16} nucleus will be dissipated before the emission of the γ -radiation, and the angular distribution measured in the laboratory frame will be the same as that in the centre of gravity frame. It is assumed, of course, that the half-life of the γ -transition is sufficiently short for any interaction between the nuclear magnetic moment and any extranuclear magnetic fields to be negligible (cf. part II). Even if the

γ -transition usually occurs before most of the O^{16} recoil momentum is destroyed, the effect of this recoil motion on the observed angular distribution will only be appreciable if there is pronounced angular correlation between the direction of emission of the α -particles and the incident proton beam. Some such correlation may, in general, be expected to exist, although in the only case in which detailed quantitative conclusions have been drawn, the angular distribution of the α -particles has, in any case, been assumed isotropic. Hence, apart from any residual effects of the primary Ne^{20*} nuclear recoil (which is an order of magnitude less than the O^{16*} recoil), the neglect of the recoil motion does not, in this particular case, involve additional assumptions.

(a) 338 and 470 keV levels

The proton bombarding energies were 350 and 480 keV in the respective cases, and the target thickness about 25 keV. No significant departure from isotropy was observed in either case (figures 10 *a*, *b*), in agreement with the results of previous investigators. For both of these resonance levels the γ -radiation is almost entirely from the 6 MeV level of O^{16} .

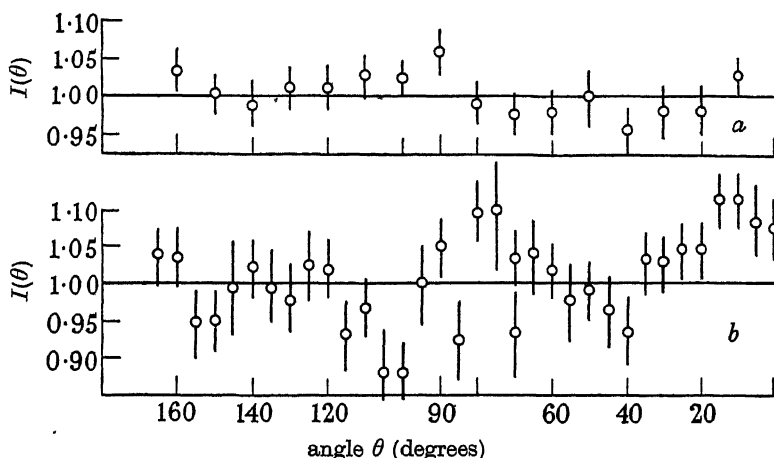
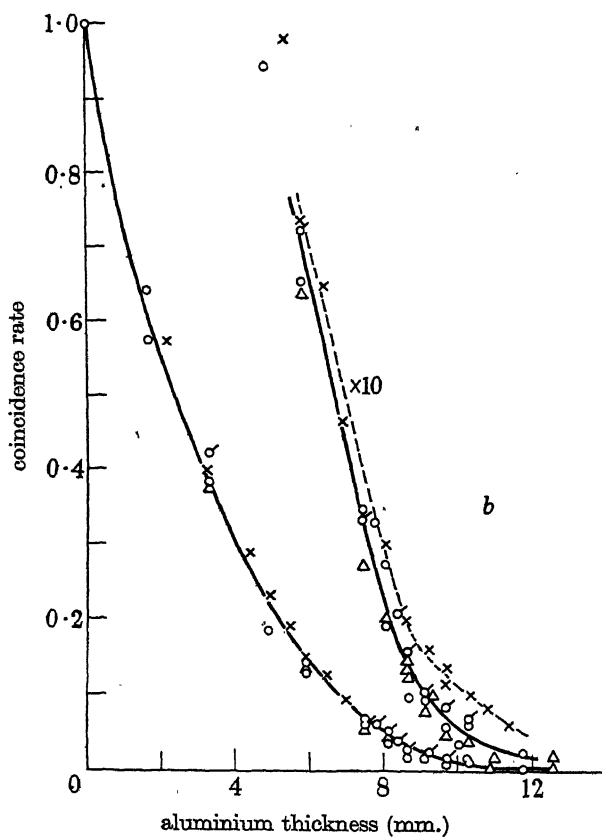
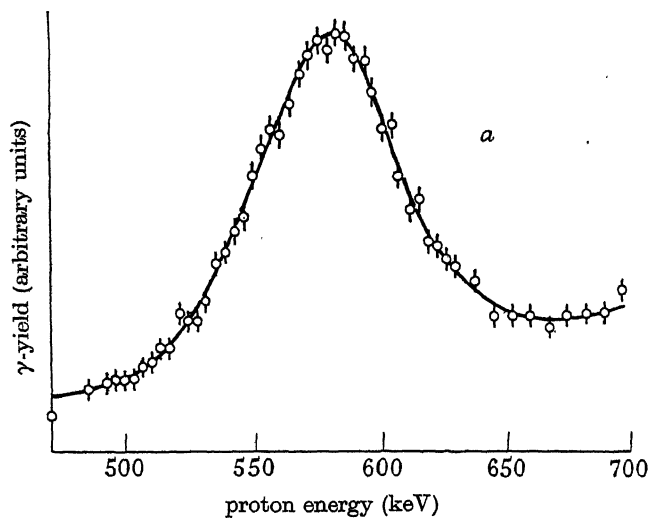


FIGURE 10. γ -radiation from reaction $F(p, \alpha) O^{16*}$; *a* 338, *b* 470 keV. $O^{16*} \rightarrow O^{16} + \gamma$.

590 keV level

The resonance at 590 keV is characterized by a much larger total width ($\Gamma \approx 40$ keV) than any of the other levels up to 1.2 MeV proton energy. Since the total width for this level is almost certainly determined primarily by the partial width, Γ_α , for the emission of an α -particle, the large observed width might be interpreted in one of the following ways:

- (i) The level is an unresolved superposition of two or more closely spaced levels.
- (ii) The radiation is from a different, lower, excited level of O^{16} , and the α -particle is emitted with correspondingly greater energy than in the case of the other, narrower resonance levels. Since the penetrability of the Coulomb potential barrier is very sensitive to the α -particle energy, a relatively small change in the latter (of the



Proton Energy: $\circ, \oslash$ 590 keV, $\times$ 670 keV, Δ 338 keV.

FIGURE 11.

order 0.8 MeV) would be sufficient to account for an increase in the level width by a factor of ten or so.

(iii) The angular momentum with which the α -particles are emitted in the break-up of the compound Ne^{20} nucleus is smaller (and the penetrability therefore larger) for this level than for any of the other levels mentioned.

In order to test the first possibility a careful check was made of the excitation function in the neighbourhood of 600 keV, using a target of thickness 1 keV and a proton energy resolution of about 2 keV. It seems safe to conclude from the results (figure 11a) that the large width is not attributable to the superposition of several narrower resonances. To test the second possibility, a measurement was made of the relative energies of the γ -radiation at 590 and 340 keV, using the method of coincident counters and absorption of the secondary Compton electrons. The results of these measurements (figure 11b) indicate that the difference in energy, at least for the major part of the γ -radiation, is less than 0.2 MeV for the two cases.

It seems likely, then, that the third explanation is the most probable.† It is possible, of course, that the level width for emission of an α -particle without potential barrier changes by as much as a factor of ten from one level to another, but this explanation appears unlikely in view of generally similar properties (apart from the width) of this and the other levels.

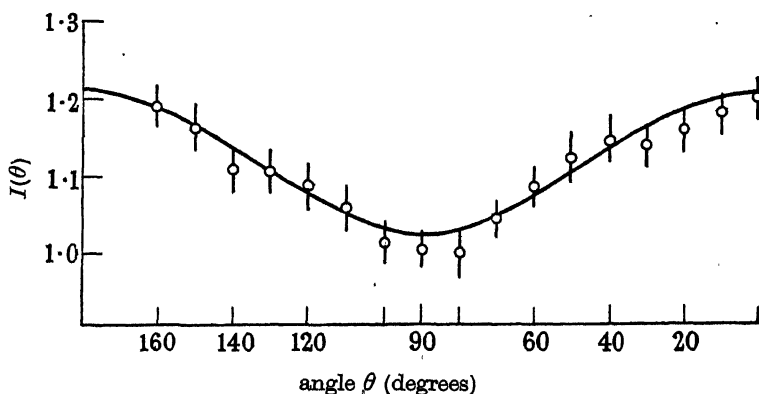


FIGURE 12. Angular distribution of γ -radiation from $\text{F}^{18}(p, \alpha)\text{O}^{16*}$. Proton energy 590 keV.

† This conclusion is supported by some recent measurements, in this laboratory, of the relative intensities of the two groups of α -particles emitted in this reaction, the main group leaving the O^{16} nucleus in the state with excitation energy of 6.1 MeV and a much weaker group, of smaller energy, corresponding to a state of O^{16} with an excitation energy 7.0 MeV. The low energy α -particle group has been observed for all of the resonance levels examined, except that at 590 keV. This is just the result one might expect, since, if for this level the main α -particle group corresponds to emission with zero angular momentum, the low energy group will be weaker by at least the ratio of the penetrabilities as calculated for zero orbital angular momentum, which would probably render it undetectable. For other levels conditions would be more favourable for the emission of the low energy particles on account of the main group being emitted with non-zero angular momentum. The intensity of this additional α -particle group is not sufficient to affect substantially the comparison of γ -energies mentioned above. (We are indebted to Dr Burcham for informing us of these unpublished measurements of the α -particle groups; cf. also Walker & McDaniel 1948).

In a quantitative examination of the level widths for this reaction Schiff (1946) has also concluded that this level at 590 keV corresponds to the emission of the α -particle with zero angular momentum, and this very plausible conclusion is further supported by relatively strong angular correlation between the directions of emission of radiation and incident proton. This was measured using a proton bombarding energy of 600 keV and a target of thickness 15 keV (figure 12). The angular distribution shows no marked asymmetry, and a satisfactory fit to the experimental points is obtained with the expression $1 + 0.2 \cos^2 \theta$ (solid line in figure 12).

(c) 665 keV level

The angular distribution of the radiation, using a bombarding energy of 675 keV and a target thickness of 15 keV, shows no significant departure from isotropy (figure 13a).

(d) 860 keV level

A proton bombarding voltage of 870 keV and a target of thickness 10 keV were used. The angular distribution (figure 13b) shows a small departure from isotropy which is probably significant. A satisfactory fit is obtained by the expression $1 + 0.1 \cos^2 \theta$. According to the measurements of Walker & McDaniel (1948), the radiation associated with this level probably contains an appreciable component (of the order 30 % of the total) with 7.0 MeV energy. The measurements reported here are not sufficient to separate the contributions of the two components.

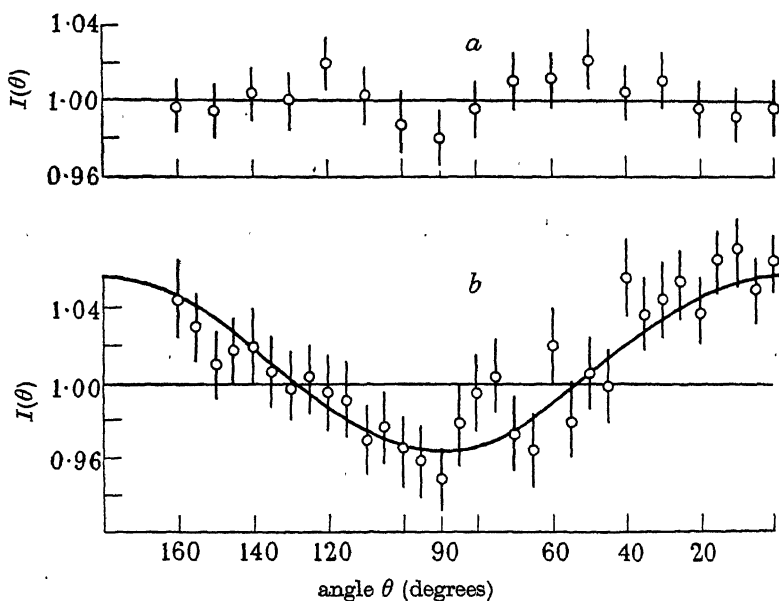


FIGURE 13. γ -radiation from $F^{19}(p, \alpha)O^{16*}$.
Proton energy *a* 675 keV; *b* 870 keV. $O^{16*} \rightarrow O^{16} + \gamma$.

One of us (M.G.N.H.) is indebted to King's College, Cambridge, and the Department of Scientific and Industrial Research for grants during the course of this work.

REFERENCES

- Ageno, M., Amaldi, E., Bocciairelli, D. & Trabacchi, G. C. 1941 *Ricerca Sci.* **12**, 139.
 Bennett, W. E., Bonner, T. W., Mandeville, C. E. & Watt, B. E. 1946 *Phys. Rev.* **70**, 882.
 Bleuler, E. & Zündt, W. 1946 *Helv. Phys. Acta*, **19**, 376.
 Bonner, T. W. & Evans, J. E. 1948 *Phys. Rev.* **73**, 666.
 Devons, S. & Hine, M. G. N. 1948 *Phys. Rev.* **74**, 976.
 Fowler, W. A., Lauritsen, C. C. & Lauritsen, T. 1948 *Rev. Mod. Phys.* **20**, 236.
 Horton, G. K. 1948 *Proc. Phys. Soc.* **60**, 457.
 Jacobs, J. A., Malmberg, P. R. & Wahl, J. S. 1948 *Phys. Rev.* **73**, 1130.
 Jaeger, J. C. & Hulme, H. R. 1935 *Proc. Roy. Soc. A*, **148**, 708.
 Schiff, L. I. 1946 *Phys. Rev.* **70**, 891.
 Van Allen, J. A. & Smith, N. M. 1941 *Phys. Rev.* **59**, 501.
 Walker, R. L. & McDaniel, B. D. 1948 *Phys. Rev.* **74**, 315.

The angular distribution of γ -radiation from light nuclei

II. Theoretical

BY S. DEVONS, PH.D. AND M. G. N. HINE,
Cavendish Laboratory, University of Cambridge

(Communicated by O. R. Frisch, F.R.S.—Received 28 January 1949—

Revised 19 April 1949—Read 16 June 1949)

A general discussion is given of the factors which determine the angular correlation between incident particles and emitted radiation, in the case of particle capture by nuclei of light atoms. The results are applied to the cases $\text{Be}^9(p, \gamma)$; $\text{C}^{12}(p, \gamma)$; $\text{C}^{13}(p, \gamma)$; $\text{F}^{19}(p, \alpha) \text{O}^{16*}$, $\text{O}^{16*} \rightarrow \text{O} + \gamma$; and in some detail to $\text{Li}^7(p, \gamma)$.

In the case of Li^7 , the variation of the angular distribution of the γ -radiation with proton energy has been studied, and the results compared with a detailed theoretical interpretation, based on the Breit-Wigner dispersion formula applied to more than one level of the compound nucleus. The comparison provides excellent support for the validity of the dispersion formula, and the conditions are very favourable for such a check.

INTRODUCTION

In the preceding paper (Devons & Hine 1949), referred to subsequently as part I, an account has been given of some measurements of the angular distribution of γ -radiation arising from the bombardment of some light nuclei with protons. In the present paper we shall indicate the general nature of the theoretical interpretation of such measurements and a more detailed interpretation of the results in one case, $\text{Li}(p, \gamma)$.

We shall consider first the angular distribution, in the centre of gravity frame of reference, of the γ -radiation resulting from the capture of particles A by target nuclei B . This radiation is to be regarded as a second-order process. One or more states of the compound nucleus, C_r , are formed by capture of A , each in some superposition or mixture of its individual substates (denoted by J_z , i.e. the z -component

of the total angular momentum J), from which radiative transitions occur to the lower states, C_0 , of this nucleus in competition with the (usually more probable) process of re-emission of the particles A , i.e. elastic scattering.

FORMATION OF THE COMPOUND NUCLEUS

The incident and bombarded nuclei have intrinsic angular momenta ('spins') $j(A), j(B)$ whose orientation with respect to the incident beam (the z -axis) is determined by $j_z(A), j_z(B)$, the z -components of the spins. A complete set of eigenstates of the system at a definite bombarding energy can be obtained by describing the z -components of the spins of the two particles separately, i.e. $j_z(A)$ and $j_z(B)$ or, equally well, by j, j_z , the magnitude of the combined angular momentum, $j(A) + j(B)$, and its z -component. In particular, the incoherent mixture, with equal weights, of all the states of the separate spins of A and B which we use to describe the randomly oriented spins in the beam and the target of an actual experiment can be replaced by a similar mixture of all the possible (j, j_z) states representing the combined spins of the system (Breit & Darling 1947). The total number of such eigenstates is, of course, the same in either representation, since

$$\sum_{|j(A)+j(B)|}^{j(A)+j(B)} (2j+1) = (2j(A)+1)(2j(B)+1).$$

The wave function representing the orbital motion of the initial state for all (j, j_z) states is the same, a plane wave in the z -direction which can be resolved into a coherent superposition of spherical waves each of which represents particles with a definite orbital angular momentum, l , whose z -component, l_z , is zero. Thus each J_z substate of C_r can be formed only from that initial (j, j_z) state which has the same z -component of angular momentum, and the incoherence of the (j, j_z) states means, therefore, that γ -radiation *intensities* from the different substates of C_r are added to obtain the observed intensity in any particular direction.

The probability amplitude for the formation of a substate of C_r can be regarded as the product of three factors.

(i) A term depending on the amplitude of the component of orbital momentum in the expansion of the initial plane wave.

(ii) The usual transformation coefficients for combination of the angular momenta (cf. Condon & Shortley 1935, chapter III), which determine the probability that the initial momenta will combine to give a state with the correct total angular momentum, $J(C_r)$.

(iii) The probability amplitudes for the formation of the actual state C_r from the component of the initial state with the same total angular momentum, $J(C_r)$. This quantity is energy dependent, includes barrier penetrability and specifically nuclear factors, and depends in general on j and l , but not on j_z , since there is no preferred direction for this particular process.

If the angular momentum conservation laws allow particles with more than one value of l to participate, the probability amplitudes for each process must be added, taking into account their relative phases, which are coherent if the same (j, j_z) states are involved.

Only orbital momenta differing by an even number of units ($\hbar$) can take part in the formation of a particular C_r , since the parity of the initial states must be the same as that of C_r and the parity of the orbital motion is even or odd with l . With charged particles of moderate energy (in the experiments described in part I the maximum proton energy was about 1 MeV) there is a large difference in barrier penetrability for difference of 2 in l , and this often means that only the component of lowest l , usually 0 or 1, makes an appreciable contribution to the process, though since components of different l are coherent, the barrier factors for amplitudes of wave functions must be compared rather than the more usual factors for intensity.

PROPERTIES OF THE γ -RADIATION

The absolute probability and multipole nature of the γ -transition from any state $C_r(J, J_z)$, of the compound nucleus to a lower state, C_0 , e.g. the ground state, depends, *inter alia*, on the angular momenta and the parities of the two states. If the transition is a 2^Δ -pole one, the quantum carries away Δ units of angular momentum (Heitler 1936), and the parity of the radiation field is related to the multipolarity as indicated in the following scheme:

Δ	odd parity	even parity
1	electric dipole	magnetic dipole
2	magnetic quadrupole	electric quadrupole
3	electric octupole	magnetic octupole

Since the wave-length of the γ -radiation (λ) is usually much larger than nuclear dimensions (R) and the probability of particular multipole transition contains the factor $(R/\lambda)^{2\Delta}$, the higher order radiation will be progressively less likely. Where conservation of angular momentum and parity permit a transition to have a mixed multipole nature, it will normally be the lowest multipole that predominates. However, since also the velocities of the nucleons in the nucleus are small compared to the velocity of light, magnetic radiation will be much less probable than electric of the same multipolarity, so that transitions of the type electric quadrupole-magnetic dipole may occur with comparable amplitudes of the two components (cf. cases of C^{12} , C^{13} below). When such mixed transitions occur, the corresponding radiation fields must be added directly and an adjustable complex parameter introduced to express the ratio of the two multipole moments. We use the expressions given in Condon & Shortley (1935, chapter IV) for the radiation fields corresponding to transitions between particular pairs of states, $C_r(J, J_z)$ and $C_0(J', J'_z)$.

The angular distribution of radiation produced by the compound nucleus formed from a particular (j, j_z) state can be found by summing, over all the states $C_r(J, J_z)$ affected, the product of the probability amplitudes for forming $C_r(J, J_z)$ and the radiation field for its decay to a state $C_0(J', J'_z)$. The magnitude squared of this sum gives the radiation intensity. The distribution found experimentally will be obtained by summing this intensity over j, j_z , and finally over J'_z , if $J' \neq 0$.

Before discussing in detail the results given in part I we shall indicate the type of angular distribution to be expected in different types of reaction.

TYPES OF DISTRIBUTION

The most marked features of an observed distribution are:

- (i) Whether isotropic or not,
- (ii) If anisotropic, whether symmetrical about the plane normal to the incident beam,
- (iii) Whether the distribution is energy dependent or not.

(i) Since the angular distribution for a transition between single J_z substates of C_r and C_0 is always anisotropic, isotropy must result from angle-dependent terms, from different substates, cancelling out. This can occur, either by some fortuitous values of the numerical constants specific to a particular reaction; or, more generally, because the distribution in the J_z substates of C_r is that of a randomly oriented assembly of nuclei, i.e. populations of equal weight and random phase. This latter is the case if $J(C_r) = 0$ or $\frac{1}{2}\hbar$, or if only the component $l = 0$ of the incident beam takes part in the process. Approximate isotropy may occur if, on account of barrier penetrability, the contribution from the $l = 0$ component predominates over that from possible higher components, $l = 2, 4, \dots$. Notice that isotropy due to an *outgoing* S -wave, a common occurrence for nucleon emission, is not possible for γ -radiation.

(ii) A transition between single states, C_r, C_0 , which have definite parities, will have an intensity distribution symmetrical to reflexion through the radiating nucleus, which, since there is always axial symmetry about the incident beam, implies back-to-front symmetry as well.

Even if several states C_r participate coherently in the transition, there will still be symmetry if all the states have the same parity.

(iii) Distributions will be energy independent if the term depending on energy in the probability for forming the states C_r can be taken as a factor outside the various summations; the total intensity will then vary, but the distribution remains constant. This will in general occur if there is only one level of the compound nucleus C_r involved, and if the proportion of states of different l in forming it remain fixed. This will be true for sharply resonant processes, and even if the width of the resonance level is comparable with the bombarding energy, although the relative contributions from different l components will generally be energy dependent, the approximate constancy of the ratio of barrier factors for different l will keep the distribution nearly independent of energy.

In practice, the spins $j(A), j(B)$ of interacting particles will be known, and total level widths for resonant reactions may give rough values for the predominant value of l . We can usefully discuss different cases, as follows:

- (a) Neither nucleus has spin; $j(A) = j(B) = 0$ and $j = j_z = 0$.
- (b) One nucleus, say A , has spin: $j = j(A)$.
- (c) Both nuclei have spin: j can take values between

$$|j(A) - j(B)| \quad \text{and} \quad |j(A) + j(B)|.$$

- (a) $l = J(C_r)$

If $l = 0$, radiation is isotropic.

If, as is usual for such reactions, $J'(C_0) = 0$, and the parities of A , B and C are all even, then the parity of C_r is even or odd according as $J(C_r)$ is even or odd, and radiation is of purely electric type. The distribution depends only on l if a single state C_r is involved.

(b) (j, j_z) substates have only one value of j , $j = j(A)$.

If $l = 0$, $J(C_r) = j$, the substates of C_r are distributed as (j, j_z) , and radiation is isotropic.

If $j \geq 1$, more than one value of l can lead to same C_r . Adjustable complex constants will be needed for the ratio of the influences of the states of different l .

(c) (j, j_z) substates can have values of j from $|j(A) - j(B)|$ to $|j(A) + j(B)|$.

Since the intensity of radiation is summed over different j , additional real adjustable constants will be needed to allow for the unknown relative efficiencies of different j states in forming C_r .

In all the cases (a), (b), (c) if several levels C_r take part (i.e. overlapping of levels occurs), extra complex parameters are needed if different levels can be formed from the same j state, and real parameters if different levels are formed from different j states.

RADIATION FROM A RESIDUAL NUCLEUS

The treatment of the angular distribution of γ -radiation from a residual nucleus, although formally similar, is usually much more involved than that of capture radiation. The process of emission of a nucleon from the compound nucleus can be regarded as formally the same as that of capture of a nucleon, but in the former case the nuclei will not generally have the normal random distribution in j_z before the emission process. One particularly simple case may occur in practice; namely, if the disintegration of the compound nucleus occurs by the emission of a particle with zero intrinsic and orbital angular momentum. In such a process the symmetry properties of the residual nucleus are identical with those of compound nucleus, and the subsequent radiation may be treated exactly as capture radiation.

In the foregoing discussion of the factors determining the angular distribution of the radiation, it has been tacitly assumed that the effect of the extranuclear structure is negligible. The possible influence of the interaction between atomic and nuclear magnetic moments on angular correlation of successive γ -radiations has been investigated theoretically by Goertzel (1946). He establishes as a conservative criterion for neglecting the interaction, in cases where the radiating nucleus is part of a normal atom, the condition

$$\tau < 0.05(\Delta\nu)^{-1},$$

where $\Delta\nu$ is the hyperfine-structure frequency interval and τ is the half-life of the nucleus in the state from which the radiation occurs. The same general condition applies in radiation processes studied here, although it is difficult to estimate the effective value of $\Delta\nu$ for a recoiling nucleus which is almost certainly not surrounded by the normal electron configuration. For capture radiation the half-life of the radiating state is determined by the combined probabilities of γ -radiation and re-emission of the incident nucleon, i.e. by the total width of the state. The total widths

of the states of the compound nuclei in all the cases of capture radiation studied here is of the order of 10^3 eV or more, corresponding to half-lives of 10^{-18} sec. or less, which is certainly far too short for the precession of the nuclear moment in any plausible local electric or magnetic field to be appreciable. For γ -radiation from a residual nucleus, on the other hand, there is no competition by particle emission, so that τ is determined solely by the probability of γ -emission, and may easily exceed the limits imposed by the above inequality. If the γ -transition has an energy of several MeV and is of low multipolarity, say not higher than quadrupole, then τ is likely to be of the order 10^{-12} sec. or less (i.e. a radiation width of 10^{-3} eV or more) and the condition above is probably satisfied.

Li⁷

The differential cross-section for the emission of a γ -quantum in the direction θ to the incident proton beam can be expressed as follows:

$$\sigma(\theta) = \sum_{j'} \sum_{j_z} \left\{ \left| \sum_r \frac{(A | H_1 | C_r) (C_r | H_2 | C_0, J'_z, \theta)}{E_p - E_r + \frac{1}{2}i\Gamma_r} \right|^2 \right\}. \quad (1)$$

In this expression A denotes the initial state, i.e. a parallel (z -direction) beam of protons and a Li⁷ nucleus in a particular (j, j_z) state, C_r the compound state of Be⁸, (C_0, J'_z, θ) the final state, i.e. the Be⁸ nucleus in the de-excited state and a γ -quantum emitted in the direction θ . H_1, H_2 represent interaction operators corresponding to the formation of Be⁸ and the emission of γ -radiation. E_p, E_r and Γ_r have the usual meanings.

Li⁷ has a total angular momentum quantum number, j , of $\frac{3}{2}$ and, we assume, odd parity, and for the ground state of Be⁸ we assume $J_0 = 0$ and even parity. From the observed angular distribution of γ -radiation near resonance (approximate isotropy), we conclude that the resonant state is formed predominantly by the $l = 0$ proton component, and therefore has odd parity. Now the cross-section for the reaction Li⁷ ($p, 2\alpha$) does not show any resonance features at an energy of 440 keV, and it is therefore legitimate to assume that the resonance state of Be⁸ formed at this energy cannot dissociate into two α -particles. Since α -particles obey Bose statistics, this will be the case for any state of Be⁸ other than those with even J and even parity. Thus the assumption that the $l = 0$ component is responsible for the resonance at 440 keV is consistent with the absence of any resonance at this energy for the Li ($p, 2\alpha$) reaction, and also with the relatively small total width of this level, which would not be consistent with the emission of energetic α -particles.

The $l = 0$ component can produce a state of Be⁸ with either $J = 2$ or $J = 1$. The γ -transition $J = 2$ odd $\rightarrow J = 0$ even, would be magnetic quadrupole, whereas $J = 1$ odd $\rightarrow J = 0$ even, would be electric dipole. In view of the large radiation width (see part I) and the fact that magnetic radiation is much less likely than electric, even for the same multipolarity, it seems very probable that the second transition is the correct one. (In this case Γ_γ for the transition is 24 eV.) With the $l = 0$ proton component the compound state $J = 1$ can be formed from the single j -state, $j = 1$, comprising three (j, j_z) components, $j_z = \pm 1, 0$. This state of the compound nucleus could also be formed from the $l = 2$ proton component, but on

account of the reduced barrier penetrability, we would expect the matrix-element corresponding to such a transition to be of the order 0.1 of that for the $l = 0$ component, and in order to keep the treatment free from undue complication we neglect this contribution.

The observed asymmetry in the distribution of the γ -radiation requires interference between at least two states of Be^8 , and the variation with energy of the cross-section indicates that, apart from the narrow level at 440 keV, the other level(s) are very wide. This is explained in a straightforward manner if it is assumed that the other state(s) are associated with the $\text{Li}^7(p, 2\alpha)$ reaction for which the variation of cross-section with proton energy is known (Heydenburg, Hudson, Inglis & Whitehead 1948) and shows no resonance maxima for proton energies up to 1 MeV. These states will have even total angular momenta, $J = 2, 4, \dots$ ($J = 0$ is not significant here since the radiative transition $J = 0 \rightarrow J = 0$ cannot occur). Analysis of the angular distribution of the α -particles in the reaction indicates that states with $J = 0, 2$ are involved in that case, and since, even if the transition $J = 4 \rightarrow J = 0$ does occur, both the proton width and the radiation width will be much smaller than for the transition $J = 2 \rightarrow J = 0$, we consider only the state(s) with $J = 2$. The transition $J = 2 \text{ even} \rightarrow J = 0 \text{ even}$, will be electric quadrupole.

A state $J = 2$ even can be produced by the l -components 1, 3, and for reasons already given we consider only the $l = 1$ component. It is irrelevant for our purpose whether we regard this $J = 2$ even state as essentially a single state of Be^8 or several broad overlapping states all with the same symmetry properties, since in either case we shall assume that the amplitude and phase of the state, or mixture of states, vary only slowly with energy. The compound state $J = 2$ even can be formed from the two possible j -states, $j = 1, j_z = 0, \pm 1$ and $j = 2, j_z = 0, \pm 1, \pm 2$.

Denoting the narrow level by the suffix n and the wide level(s) by w we can write, from (1):

$$\begin{aligned} \sigma(\theta) &= \pi\lambda^2 \sum_{j, j_z} \left| \frac{(A(l=0; j=1) | H_1 | C_n) (C | H_2 | C_0, \theta)}{[n]} \right. \\ &\quad \left. + \frac{(A(l=1; j=1, 2) | H_1 | C_w) (C_w | H_2 | C_0, \theta)}{[w]} \right|^2 \quad (2) \\ &= \pi\lambda^2 \left\{ \sum_{j_z=0, \pm 1} \left| \frac{(A(0, 1, j_z) | H_1 | C_n(j_z)) (C_n(j_z) | H_2 | C_0)}{[n]} \right. \right. \\ &\quad \left. + \frac{(A(1, 1, j_z) | H_1 | C_w(j_z)) (C_w(j_z) | H_2 | C_0)}{[w]} \right|^2 \\ &\quad \left. + \sum_{j_z=0, \pm 1, \pm 2} \left| \frac{(A(1, 2, j_z) | H_1 | C_w(j_z)) (C_w(j_z) | H_2 | C_0)}{[w]} \right|^2 \right\}, \quad (3) \end{aligned}$$

where $A(0, 1, j_z)$ indicates an initial state with $l = 0, j = 1, j_z$, etc., and the state C is designated $C(j_z)$ to denote the fact that $J_z = j_z$ on account of the choice of the z -axis. $[n], [w]$ are abbreviations for the complex quantities $(E_p - E_n + \frac{1}{2}i\Gamma_n)$ and $(E_p - E_w + \frac{1}{2}i\Gamma_w)$.

The matrix elements corresponding to the formation of the compound nucleus can be factorized into two terms, the first resulting from the transformation from a l, l_z, j, j_z ($l_z = 0$) representation to one in which J, J_z assume definite values, and the second factor denoting the extent to which H_1 couples the initial state of appropriate symmetry with the actual states of the compound nucleus (C_n and C_w). Using the usual transformation coefficients for the first factor we can write

$$(A(0, 1, j_z) | H_1 | C_n(j_z)) = f_n(E_p), \quad (4a)$$

(the first factor is unity in this case).

$$\left(\begin{matrix} & +1 \\ A(1, 1, & 0 \\ & -1 \end{matrix} \middle| H_1 \middle| C_w \begin{pmatrix} +1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} \\ \sqrt{\frac{2}{3}} \\ 1/\sqrt{2} \end{pmatrix} \cos \chi e^{i\alpha} f_w(E_p) \quad (4b)$$

and

$$\left(\begin{matrix} & +2 \\ & +1 \\ A(1, 2, & 0 \\ & -1 \\ & -2 \end{matrix} \middle| H_1 \middle| C_w \begin{pmatrix} +2 \\ +1 \\ 0 \\ -1 \\ -2 \end{pmatrix} = \begin{pmatrix} -\sqrt{\frac{2}{3}} \\ -\sqrt{\frac{1}{6}} \\ 0 \\ +\sqrt{\frac{1}{6}} \\ +\sqrt{\frac{2}{3}} \end{pmatrix} \sin \chi f_w(E_p). \quad (4c)$$

$f_n(E_p), f_w(E_p)$ are complex functions of E_p , assumed to be slowly varying compared with $[n]^{-1}$. α is a phase angle introduced to indicate the unknown phase difference between the matrix elements represented by (4a) and (4b). The factors $\cos \chi, \sin \chi$ represent the relative coupling by H_1 of the relevant components of the initial states, $l = 1, j = 1$ and $l = 1, j = 2$ with the state C_w . From (3) it is clear that the phase of (4c) is irrelevant.

For the matrix elements $(C_n(j_z) | H_2 | C_0)$ and $(C_w(j_z) | H_2 | C_0)$ corresponding to electric dipole and quadrupole transitions respectively, we use the following expressions for the electromagnetic fields (E) at unit distance from the radiating nucleus in the direction (θ, ϕ) (Condon & Shortley 1935).

$$E_{\text{dipole}} = \left\{ \begin{array}{ll} -\sqrt{\frac{1}{2}} K^2 D(\theta_0 \cos \theta - i\phi_0) e^{-i\phi} & \begin{matrix} \text{transition} \\ j_z \rightarrow j_z \\ 1 \rightarrow 0 \end{matrix} \\ -K^2 D(\theta_0 \sin \theta) & 0 \rightarrow 0 \\ +\sqrt{\frac{1}{2}} K^2 D(\theta_0 \cos \theta + i\phi_0) e^{i\phi} & -1 \rightarrow 0 \end{array} \right\} \quad (5a)$$

$$E_{\text{quad.}} = \left\{ \begin{array}{ll} \frac{1}{2} i K^3 Q \sin \theta e^{\mp 2i\phi} (\theta_0 \cos \theta \mp i\phi_0) & \pm 2 \rightarrow 0 \\ \mp \frac{1}{2} i K^3 Q e^{\mp i\phi} (\theta_0 \cos 2\theta \mp i\phi_0 \cos \theta) & \pm 1 \rightarrow 0 \\ -\frac{1}{2} \sqrt{3} i K^3 Q \sin \theta \cos \theta \theta_0 & 0 \rightarrow 0 \end{array} \right\}. \quad (5b)$$

In these expressions D, Q are complex quantities representing the strengths of the dipole and quadrupole moments, K is the wave number of the radiation, and θ_0, ϕ_0 unit vectors in the directions of increasing θ and ϕ . Using these expressions for the electric field, taking EE^* as a measure of the intensity of radiation in the direction (θ, ϕ) , and neglecting constant factors which do not affect the angular distribution,

we obtain from (3), (4) and (5) the following expression for the variation of intensity with angle:

$$\begin{aligned}
 I(\theta) = & \left| \frac{A}{\sqrt{2}} (\theta_0 \cos \theta - i\phi_0) e^{i\phi} + \frac{B}{2\sqrt{2}} e^{-i\phi} (\theta_0 \cos 2\theta - i\phi_0 \cos \theta) \right|^2 \\
 & + \left| \frac{A}{\sqrt{2}} (\theta_0 \cos \theta + i\phi_0) e^{i\phi} + \frac{B}{2\sqrt{2}} e^{i\phi} (\theta_0 \cos 2\theta + i\phi_0 \cos \theta) \right|^2 \\
 & + |A\theta_0 \sin \theta + B\theta_0 \sin \theta \cos \theta|^2 \\
 & + \left| \frac{C}{\sqrt{6}} \sin \theta e^{-2i\phi} (\theta_0 \cos \theta - i\phi_0) \right|^2 + \left| \frac{C}{\sqrt{6}} \sin \theta e^{2i\phi} (\theta_0 \cos \theta + i\phi_0) \right|^2 \\
 & + \left| \frac{C}{\sqrt{24}} e^{-i\phi} (\theta_0 \cos 2\theta - i\phi_0 \cos \theta) \right|^2 + \left| \frac{C}{\sqrt{24}} e^{i\phi} (\theta_0 \cos 2\theta + i\phi_0 \cos \theta) \right|^2,
 \end{aligned} \tag{6}$$

where we use the abbreviations

$$\begin{aligned}
 A &= \frac{K^2 |f_n(E) D| e^{i\rho}}{[n]}, \quad B = \frac{\frac{1}{2} K^3 \cos \chi e^{i\alpha} |f_w(E) Q| e^{i\tau}}{[w]} \\
 C &= \frac{\frac{1}{2} K^3 \sin \chi |f_w(E) Q| e^{i\tau}}{[w]}.
 \end{aligned} \tag{7}$$

Expression (6) simplifies to

$$\begin{aligned}
 I(\theta) &= 2AA^* + \frac{1}{4}BB^*(1 + \cos^2 \theta) + (AB^* + B^*A) \cos \theta + CC^*(\frac{1}{8} + \frac{1}{4} \sin^2 \theta), \\
 \text{i.e. } I(\theta) &= 2|A|^2 + \frac{|B|^2}{4}(1 + \frac{5}{3}t^2) + \frac{|B|^2}{4}(1 - t^2) \cos^2 \theta + 2|A||B| \cos \theta \cos \delta,
 \end{aligned} \tag{8}$$

where

$$\begin{aligned}
 t &= \tan \chi, \quad \delta = (\rho - \alpha - \tau - \gamma_n + \gamma_w), \quad \tan \gamma_n = \Gamma_n/2(E - E_n), \\
 \tan \gamma_w &= \Gamma_w/2(E - E_w).
 \end{aligned}$$

In order to compare the theoretical expression (8) with the experimental results we need to know the following energy-dependent quantities: $|A|/|B|$, δ and t . We make the assumptions that the variation of t with energy is negligible and that of δ is due to the variation of γ_n and γ_w . γ_w varies only slightly with energy, and we have computed (8) first on the basis of ignoring this variation and secondly by assuming that the wide level of Be⁸ has a broad resonance ($\Gamma_w \sim 1$ MeV) at an energy E_w of 3 MeV (Heydenburg *et al.* 1948). The variation of γ_n is obtained by using the values, $\frac{1}{2}\Gamma_n = 6$ keV and $E_n = 440$ keV (Bonner & Evans 1948). Similarly, in estimating the variation of $|A|/|B|$ with energy we have taken χ , D and Q to be energy independent, made the two corresponding assumptions for the variation of $[w]$, and for the variation of the ratio $f_n(E)/f_w(E)$ with energy we have used the calculations of Christy & Latter (1948), of the relative penetrabilities of the Coulomb potential barrier for S -wave and P -wave protons. Despite these numerous simplifications we do not consider it likely that our assumptions will lead to substantial inaccuracies, since the predominant variation in $I(\theta)$ will arise from the variation of γ_n , $[n]$, both of which can be evaluated from experimental data with fair accuracy.

A more serious difficulty in interpreting the actual variation of γ -intensity with

angle is the fact that the radiation consists of two components, about 17 and 14 MeV, and the quantities D , Q and δ will differ for the two γ -transitions. We have estimated the *relative* variation of D and Q with proton energy from our own results (part I, figure 7) and from the more accurate recent measurements of Walker & McDaniel (1948). In the absence of other information we have assumed δ to be the same for both γ -transitions, an assumption which could, in principle, be tested experimentally by examining the angular distribution of the two γ -components separately. The actual observed distribution will then be the sum of two expressions like (8), allowance being made for the different ratio D/Q in the two cases and the different efficiency of detection of γ -radiation of the two energies. This assumption gives for the expected angular distribution

$$I(\theta) = 2.3 |A|^2 + \frac{|B|^2}{2} (1 + \frac{5}{3}t^2) + \frac{|B|^2}{2} (1 - t^2) \cos^2 \theta + 3.23 |A| |B| \cos \delta \cos \theta, \quad (9)$$

where $|A|$ and $|B|$ still refer to the 17 MeV component as in (7). The quantity t can be estimated from the term in $\cos^2 \theta$ in the angular distribution at proton energy of 890 keV. At this energy $|A|/|B| \ll 1$, so that we can neglect the $|A|^2$ term and thereby obtain t directly. Comparison with experiment gives $t^2 = 0.6$ (part I, figure 6). We next use the result that the maximum asymmetry occurs, according to (8), when $|B|/|A| = 1.38$. The experimental results (part I, figure 4) allow us to estimate the quantity $\frac{f_w(E)}{f_n(E)} \frac{D}{Q}$ and hence $\left| \frac{B}{A} \right|$ at any particular energy. Lastly we wish to know δ . This may be obtained directly from the experimental results by noticing the energy at which the asymmetry is zero, i.e. $\cos \delta = 0$, but this procedure is unreliable for two reasons. In the first place the direct experimental determination of δ presents difficulty on account of the finite target thickness, and the indirect deduction of δ from the thick target measurements requires an accurate knowledge of the variation of cross-section with energy. Secondly δ may be different for the two γ -components and an estimate of the effective average to be used in (8) cannot be obtained reliably from measurements near resonance. We have therefore chosen a value of δ which is a compromise between the directly observed value and that required to give the measured maximum asymmetry (part I, figure 4). It should also be pointed out that any slight non-uniformity in the targets used, and particularly even *small* local regions where the target was substantially thicker than average, would not be revealed by a substantial modification of the total cross-section curve, but might result in large reductions in the asymmetry below the values that would be obtained with an ideally thin target. The small discrepancies between the results shown in part I, figures 4 and 5, are probably attributable to this source.

The curve *a* in figure 1 shows the theoretically predicted variation of $I(15^\circ)/I(145^\circ)$ using the values

$$t^2 = 0.6, \quad \left| \frac{B}{A} \right| = 0.065, \text{ at } 440 \text{ keV}, \quad E_n = 440 \text{ keV}, \quad \Gamma_n = 12 \text{ keV},$$

$$E_w = 3.0 \text{ MeV}, \quad \Gamma_w = 1.0 \text{ MeV}, \quad \delta = \frac{1}{2}\pi \text{ at } 428 \text{ keV}.$$

Curve *b* is the theoretical curve obtained if the variation of $[w]$ with energy (with respect to both magnitude and phase) is ignored and the quantity replaced by its

average value. It can be seen from figure 1 that although the agreement between theory and experiment is not exact, it is as good as could be expected in view of the theoretical approximations and the experimental limitations. Two features are, however, quite clear. First the variation in $|A|$ at high proton energies (600 to 1000 keV), where $E - E_n \gg \Gamma_n$ is exactly that to be expected from the term $(E - E_n + \frac{1}{2}i\Gamma_n)^{-1}$ and is independent of any exact knowledge of Γ_n . Secondly the

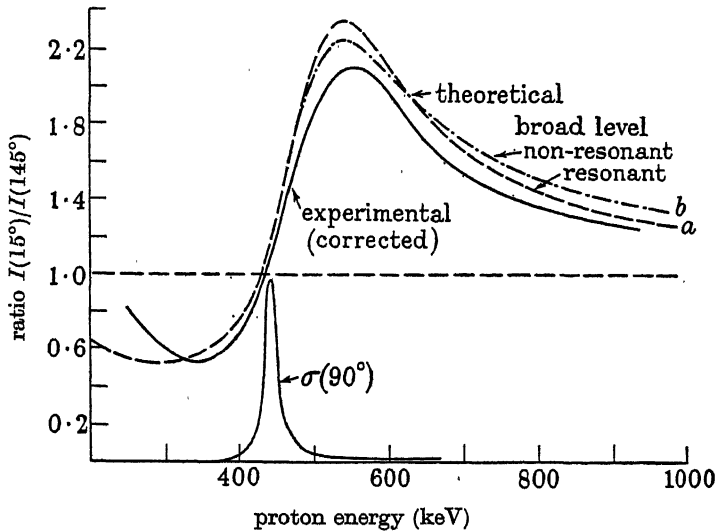


FIGURE 1. $\text{Li}^7(p, \gamma)$. Target thickness $\begin{cases} 12 \text{ keV} \\ 20 \text{ keV} \end{cases}$ and 6 keV.

variation in the asymmetry near resonance clearly indicates the change in phase as E passes through the value E_n arising from the term $\frac{1}{2}i\Gamma_n$. The interference between the two levels, and in particular the large value of $|A/B|$ at resonance and the approximate constancy of $|B|$ enable us to study the shape of the resonance term, $|A|$, at an energy some hundred half-widths from the resonance peak.

The results fully confirm the correctness in this case of the Breit-Wigner dispersion formula on which equation (1) is based. In the following cases we shall be dealing with essentially simpler cases in which only one level of the compound nucleus is involved and we shall give only the results of the calculations.

Be⁹

The small departure of the observed angular distribution from isotropy, $(1 + 0.09 \sin^2 \theta)$, is most readily interpreted as due to the contribution of the $l = 2$ component of the incident proton beam to a process which proceeds primarily through the $l = 0$ component. The resonance level of B^{10} , at a proton energy of 990 keV would then be a $J = 2$ or 1, odd state (assuming that the ground state of Be^9 is a $j = \frac{3}{2}$ odd state); and the lowest multipole radiation for the γ -transition to the ground state of B^{10} ($j = 3$ even, assumed) would be either electric dipole ($J = 2$), or electric octupole and magnetic quadrupole mixed ($J = 1$). The γ -radiation widths, as deduced from the observed cross-section, would be 20 and 33 eV respectively.

The former value is reasonable for electric dipole radiation of energy 7.4 MeV, but the latter is improbably large for radiation of such high multipolarity.

In both cases there would be some coherent interference between the contributions from the $l = 0$ and $l = 2$ components. For $J = 2$ the responsible j -state would be $j = 2$ (proton and Be^9 spins parallel) and for $J = 1$, the $j = 1$ state (spins anti-parallel). One would expect then the order of magnitude of the term in $\cos^2 \theta$ in the angular distribution to be proportional to the ratio of the matrix elements for formation of the compound nuclear state by the two l -components. On account of the difference in barrier penetrability for S and D proton waves, one would expect the matrix elements for the latter to be smaller than for the former by a factor of about 0.15 (Christy & Latter 1948), which is of the correct order of magnitude to explain the experimental results.

The presence or otherwise of terms of the type $\cos^4 \theta$ in the angular distribution would be decisive in determining whether the radiation is dipole, but in view of the smallness of the contribution from the D -wave, and the fact that there is coherent interference between S and D contributions, the term in $\cos^4 \theta$ is likely to be, even if it is present, an order of magnitude smaller than the $\cos^2 \theta$ term, and therefore not excluded by the experimental results. The most plausible assignment for the resonance level is then, $J = 2$, odd, and all the data are consistent with this.

C^{12}

Both the large total width (35 keV) and the isotropic distribution of the radiation suggest that the $l = 0$ proton component is responsible. Assuming the C^{12} ground state is a $j = 0$ even state, this would mean that the compound N^{13} nucleus is formed in a $J = \frac{1}{2}$ even state, and no contribution from proton components of larger l would be possible. If N^{13} ground state is $j = \frac{1}{2}$ odd, then the radiation could be electric dipole which is consistent with the γ -radiation width, 0.6 eV; remembering that the γ -transition energy is only 2.3 MeV, and the radiation width is strongly dependent on γ -energy (approximately as the cube of the energy for dipole radiation).

The analysis of the angular distribution is considerably simplified by the fact that the bombarded nucleus has zero angular momentum and even parity, since as a result, the compound state can be formed with only two values of J_z , $\pm \frac{1}{2}$, and these are formed with equal weight and incoherently ($j_z = \pm \frac{1}{2}$). Thus the angular distribution is determined by the value of J , the total momentum of the compound state, and does not depend on l explicitly. Only one particular l -component is effective (i.e. for a single level of the compound nucleus), and states will be even or odd according as l is even or odd. Thus in the case of transitions between the resonance level and ground state of N^{13} we have the following possibilities:

l	0	1	1	2	2	3
J	$\frac{1}{2}$		$\frac{3}{2}$		$\frac{5}{2}$	
parity	even	odd	odd	even	even	odd
lowest radiation multipole	elect. dip.	mag. dip.	mag. dip. + elect. quad.	elect. quad.	mag. quad. + elect. octupole	elect. quad.

The corresponding angular distributions are:

$$J = \frac{1}{2}, \quad \text{isotropic (a general result),}$$

$$J = \frac{3}{2} \text{ even} \quad f(\theta) \sim 1 - \frac{2}{5} \cos^2 \theta,$$

$$J = \frac{3}{2} \text{ odd}, \quad f(\theta) \sim 1 + \left[\frac{1 - 3x^2 - 6x \sin \xi}{1 + 5x^2 + 2x \sin \xi} \right] \cos^2 \theta,$$

$$J = \frac{5}{2} \text{ odd}, \quad f(\theta) \sim 1 + 6 \cos^2 \theta - 5 \cos^4 \theta.$$

(x is proportional to the ratio of magnetic dipole to electric quadrupole moments and ξ the phase difference between the two radiation fields in a mixed transition.)

The observed angular distribution for $\text{C}^{13}(p, \gamma)$ is consistent with $J = \frac{1}{2}, l = 0$ or 1, or possibly, but very unlikely, with particular values for x and ξ resulting in very small anisotropy in the case of a $J = \frac{3}{2}$ odd state.

C^{13}

Two γ -components contribute to the observed angular distribution, the 8.1 and 5.8 MeV components, the former being a direct transition from the resonance level to the ground state of N^{14} . Since our experimental arrangement was about twice as sensitive to the radiation of the higher energy, and in the absence of reliable data about the intermediate states in N^{14} , we shall consider the direct transition only.

The observed level width and isotropy both indicate that the $l = 0$ proton component is mainly responsible but this does not determine J uniquely. The possibilities, assuming the ground state of C^{13} is $j = \frac{1}{2}$ odd, and of N^{14} , $j = 1$ even, are:

l	1	0	1	0, 2	1, 3	2	3	2, 4
$j = j(A) + j(B)$	1	0	0 and 1	1	1	0 and 1	0 and 1	1
J	0		1		2		3	
parity	even	odd	even	odd	even	odd	even	odd
lowest radiation multipole	mag. dip.	elect. dip.	mag. dip. + elect. quad.	elect. dip.	mag. dip. + elect. quad.	elect. dip.	elect. quad.	—

The corresponding angular distributions are:

$$J = 0, \quad \text{isotropic, a general result.}$$

$$J = 1 \text{ even}, \quad 1 + \left\{ \frac{(1 + x^2 - 6x \sin \xi)(2t - 1)}{(1 + x^2)(3 + 2t) + 2x(2t - 1) \sin \xi} \right\} \cos^2 \theta.$$

$$J = 1 \text{ odd}, \quad \text{isotropic (if } l = 2 \text{ component is neglected).}$$

$$J = 2 \text{ even}, \quad 1 + \left\{ \frac{\frac{3}{7} - \frac{3}{4}x^2 - 3x \sin \xi}{1 + \frac{4}{8}x^2 + x \sin \xi} \right\} \cos^2 \theta.$$

$$J = 2 \text{ odd}, \quad 1 - \frac{3(1 + 2t)}{9 + 10t} \cos^2 \theta.$$

$$J = 3 \text{ even}, \quad 1 + \frac{4(1 + 3t)}{(3 + 2t)} \cos^2 \theta - \frac{5(1 + 6t)}{3(3 + 2t)} \cos^4 \theta.$$

x, ξ have the same significance as in the previous example (C^{12}); t (real and positive) is a factor denoting the relative contributions from the $j = 0$ and $j = 1$ spin-states to the formation of the compound nucleus. The most likely possibilities are $J = 0$ and $J = 1$, although there is, as mentioned above, a chance that the values of x, ξ reduce the coefficient of the $\cos^2 \theta$ term in a particular case to a value too small to give detectable anisotropy.

If we take into account the possible contribution from the $l = 2$ component to the $J = 1$ odd state, the angular distribution becomes

$$1 + \frac{3\beta^2 - 6\beta \cos \eta}{4 - 7\beta^2 + 2\beta \cos \eta} \cos^2 \theta,$$

where $\beta e^{i\eta}$ is the ratio of the matrix elements for the $l = 2$ and $l = 0$ components. (As in the case of Be^9 there is coherence between the contributions arising from the $l = 2$ and $l = 0$ components since the $J = 1$, odd, state can be produced from the $j = 1$ spin-state by both l -components.) From the ratio of penetrabilities of S and D proton waves we would expect β to be about 0.085, so that unless $\eta \sim \frac{1}{2}\pi$ there would be detectable anisotropy for this value of J .

The measured cross-section for the process corresponds to a radiation width of 30 eV if $J = 0$, and 10 eV if $J = 1$ (assuming that half the observed radiation is attributable to the 8.1 MeV transition), both of which are consistent with electric dipole radiation.

F^{19}

In attempting to interpret the results for the reaction $\text{F}^{19}(p, \alpha) \text{O}^{16*}$ we must take into account the fact that none of the resonance levels for the emission of a low-energy α -particle and the production of O^{16*} (i.e. the γ -radiating state of O^{16} at about 6 MeV) can emit α -particles with the full energy available, leaving O^{16} in the ground state.† We can account for this fact in a straightforward manner by assuming that these radiation-producing resonance levels of Ne^{20} have either even parity and odd angular momentum or odd parity and even angular momentum, symmetry properties which would prohibit the dissociation into normal O^{16} and an α -particle. Assuming the ground state of F^{19} is $j = \frac{1}{2}$ odd, we then have the following possibilities for l -components up to 3.

l	$j = j(A) + j(B)$	$J(\text{Ne}^{20})$	parity
0	0	0	odd
2	0, 1	2	odd
1	0, 1	1	even
3	0, 1	3	even

To some extent the value of l can be estimated from the size of the cross-section at resonance; and the orbital angular momentum with which the α -particle is emitted from the compound nucleus from the measured total width of the resonance levels (Schiff 1946; Bonner & Evans 1948). Such an analysis indicates, although with no

† The excitation function for energetic α -particles from the reaction $\text{F}^{19}(p, \alpha) \text{O}^{16}$ show no resonance levels at the energies for which the γ -radiation resonances are observed.

great certainty, the following values for l , the angular momentum of the incident protons, and L , that of the emitted α -particles:

proton energy	$l(p)$	$L(\alpha)$	Ne ²⁰ state
338	0	1	0 odd
470	2, 1	2	2 odd or 1 even
590	2, 1	0	2 odd or 1 even
665	2, 1	2	probably 1 even
870	2, 1	2, 3	—

338 keV level. Isotropy of the γ -radiation (and also of the low energy α -particles) will be perfect if the $l = 0$ proton component is responsible, since the selection rule for prohibition of emission of energetic α -particles limits the Ne²⁰ state to a $J = 0$ odd, and therefore no other l -component will contribute. The direct γ -transition to the ground state of Ne²⁰ cannot take place from this level in accord with observation (Devons & Hereward 1948). Emission of an α -particle with orbital momentum 1 will result in a $j = 1$ even, state of O^{16*}.

470 and 665 keV levels. If the first assignments of angular momenta given in the above table, $l = 2$, $L = 2$, are correct then one might expect some departure from isotropy in the γ -radiation associated with both these levels, but the anisotropy is likely to be small, possibly too small to be revealed with the limited experimental accuracy available. However, the assumption of the value $l = 2$, and therefore $J = 2$, odd, for the Ne²⁰ level, leads to difficulties in explaining the observed 12 MeV γ -transition from the 665 keV level with a γ -radiation width of about 2 eV, a value which is rather large for magnetic quadrupole radiation. The assignment $L = 2$ to the orbital momentum of the α -particle also leads to difficulties, since it corresponds to the production of the 6 MeV excited O^{16*} nucleus in an odd state, whereas the evidence from the 338 keV level indicates even parity for this state.

A more likely assignment to the 665 keV level is $l = 1$ for the proton component and $J = 1$, even, for the Ne²⁰ state, magnetic dipole radiation for the 12 MeV transition to the ground state of Ne²⁰, and $L = 2$ for the orbital momentum of the α -particle. It still remains to be explained why the break-up of the compound Ne²⁰ nucleus into O^{16*} and an α -particle does not take place by emission of an α -particle with $L = 0$. That this does not, in fact, happen is indicated by lack of detectable anisotropy in the γ -radiation. The assignment of $J = 2$ odd, to the resonance level of Ne²⁰ at 470 keV proton energy would be consistent with the non-observation of the 12 MeV γ -radiation from this level.

590 keV level. We have already indicated reasons for believing that this level is associated with the emission of α -particles with zero orbital momentum, and therefore the symmetry properties of the resonance level of Ne²⁰ and the γ -emitting O^{16*} state are identical. Possible values for momentum and parity are therefore 1 even; 2 odd; 3 even; etc. The magnitude of the cross-section at resonance indicates formation of the compound nucleus by the proton component $l = 1$ or 2, and to obtain consistency with the assignments already made it is necessary to choose $l = 1$, and the radiation from O^{16*} (1 even), is then magnetic dipole.

The compound Ne²⁰ state can be formed from both spin combinations, $j = 1$, and $j = 0$, together with the $l = 1$ proton component. Denoting by t the ratio of the

coupling between each of these two initial states ($j = 1, l = 1$) and ($j = 0, l = 1$), and the resonance state of Ne^{20} , we obtain for the angular distribution of the γ -radiation from O^{16} :

$$f(\theta) = 1 + \left(\frac{t-2}{t+2} \right) \cos^2 \theta.$$

Comparison with the experimental results (part I, figure 12) indicates a value of 3 for t . As in the case of $\text{Li}^7(p, \gamma)$ we see that there is no strong tendency towards conservation of intrinsic angular momentum, but this does not necessarily imply that spin momentum itself is not approximately conserved in reactions of this type, since the intrinsic momenta of the complex nuclei involved ($\text{Li}^7, \text{F}^{19}$) cannot be assumed to be entirely spin momentum.

870 keV level. It is not possible from the measurements available to decide between the several possible assignments for this level, since, apart from the small anisotropy and limited accuracy, there is an additional complication due to an appreciable contribution from the γ -transition from another state of O^{16} , with excitation energy at 7.0 MeV. The experimental technique employed in the present measurements was insufficiently refined to separate the contributions of energy 6.1 and 7.0 MeV.

One of us (M.G.N.H.) is indebted to King's College, Cambridge, and the Department of Scientific and Industrial Research for grants during the course of this work.

REFERENCES

- Bonner, T. W. & Evans, J. E. 1948 *Phys. Rev.* **73**, 666.
 Breit, G. & Darling, B. T. 1947 *Phys. Rev.* **71**, 402.
 Christy, R. F. & Latter, R. 1948 *Rev. Mod. Phys.* **20**, 185.
 Condon, E. V. & Shortley, G. H. 1935 *Theory of atomic spectra*. Cambridge University Press.
 Devons, S. & Hereward, H. G. 1948 *Nature*, **162**, 331.
 Devons, S. & Hine, M. G. N. 1949 *Proc. Roy. Soc. A*, **199**, 56.
 Goertzel, G. 1946 *Phys. Rev.* **70**, 897.
 Heitler, W. 1936 *Proc. Camb. Phil. Soc.* **32**, 112.
 Heydenburg, N. P., Hudson, C. M., Inglis, D. R. & Whitehead, W. D. 1948 *Phys. Rev.* **74**, 405.
 Schiff, L. I. 1946 *Phys. Rev.* **70**, 891.
 Walker, R. L. & McDaniel, B. D. 1948 *Phys. Rev.* **74**, 315.

Spectroscopic studies of low-pressure flames. II. Effective translational and rotational temperatures from CH bands

BY A. G. GAYDON AND H. G. WOLFHARD

Chemical Engineering Dept, Imperial College, London

(Communicated by Sir Alfred Egerton, F.R.S.—Received 2 February 1949)

[Plate 1]

Using a Fabry-Perot interferometer, the Doppler widths of lines of the CH bands emitted from the reaction zone of an oxy-acetylene flame have been examined. The experimental arrangement and method of computation are described. The effective translational temperatures of the excited CH radicals are around 4000°K in the low-pressure flames, compared with theoretical maximum flame temperatures around 2700°K . At atmospheric pressure the translational temperature is near the expected flame temperature. The effective rotational temperatures have also been measured for the CH bands 4315 and $3900\text{\AA}$ and do not differ much from the theoretical maximum flame temperatures for any of the several flames examined. The method of excitation of the CH is discussed. It seems likely that the radicals are excited by collision with other active species in the flame; the high translational temperature shows that we do not have normal thermal excitation, but there are also difficulties in attributing the effects to true chemiluminescence. For the weak CH band at $3143\text{\AA}$ the effective rotational temperature is very high, especially at low pressure.

INTRODUCTION

By spectroscopic methods it is possible to learn something about the distribution of energy among its various possible forms in the reaction zone of a flame, where thermal equilibrium is not to be expected. Also it is of importance to find out to what extent the light emission in flames is due to thermal causes and to what extent it is chemiluminescence. This knowledge is necessary for a reliable application of spectroscopic observations to combustion problems, and also may give us detailed information about collision processes in reacting gases which is of wider application.

In a previous paper (Gaydon & Wolfhard 1948*a*) we have stressed the advantage of using flames at very low pressure (pressures down to 1 mm. of mercury) to give a very thick reaction zone and a flat flame which can be examined in detail, and have reported on temperature measurements on acetylene flames by thermocouples, by the spectrum-line reversal method, and from studies of the rotational temperature of the excited OH radicals. It was then shown that the OH emission from hydrocarbon flames was mainly due to a chemiluminescence, which may be due to reaction of CH radicals with O_2 molecules to form CO and excited OH at high rotational temperature. We have also given other reports on this work, especially relating to the OH radical (Wolfhard 1939; Klaukens & Wolfhard 1948; Gaydon & Wolfhard 1948*c*). In the present paper we propose to concentrate on results for the emission from the CH radical, and in particular to describe the experimental method and results of measuring the translational temperature of the excited CH radicals by the Doppler broadening of the spectrum lines. While Doppler shifts are frequently used to determine stellar velocities and occasionally gas velocities, and pressure broadening has been used as a measure of gas pressure in flame gases (Bundy &

Strong 1948), this is probably the first detailed account of the use of Doppler broadening for the measurement of temperatures in flames. The strong bands of CH around 4315 and 3900 Å are among the strongest features of the emission spectrum of the inner cones of premixed flames of hydrocarbons and many organic compounds, but they do not fit readily into most detailed theories of combustion processes, as do the OH radicals, and it would be of interest if we could settle whether the strength of these bands resulted from the high concentration of CH radicals in the flames or was merely the result of chemiluminescence in some side reaction.

TRANSLATIONAL TEMPERATURES; EXPERIMENTAL

Theory

The width of a spectrum line depends on its natural width, resulting from the uncertainty principle, collision broadening, due to interference by neighbouring inert molecules, to resonance broadening due to interference by similar molecules, and to Doppler broadening due to the random thermal motions of the emitting molecules. To make use of the Doppler broadening to determine the effective translational temperatures of the emitting molecules, we must be sure that broadening from the other causes is negligible compared with the Doppler broadening. The natural breadth of a spectrum line is extremely small and may safely be neglected in comparison with Doppler broadening above room temperature. The resonance pressure broadening is most serious in the case of atomic spectra, especially for the resonance lines and certainly precludes the use of any strong atomic lines such as the sodium *D* lines, and for other atomic lines it would be necessary to ensure a low partial pressure of the atomic species; for molecular spectra, however, we can again safely neglect resonance broadening. The normal collision broadening will, of course, be a function of pressure and will affect all molecules to some extent at high gas pressure; we may say at this stage that the breadth of the CH lines in a flame at atmospheric pressure was close to that expected for Doppler broadening alone, so that in the flames at low pressure (from 1/50 to 1/500 atm.) the pressure broadening must certainly be quite negligible. Thus for the spectra of molecules such as CH, OH, C₂ and NH it should be possible to obtain information about the effective translational energy of the molecules from a study of line breadth.

It can be shown (see Tolansky 1947) that the half-breadth, *b*, of a spectrum line due to Doppler broadening is given by

$$b = 2\sqrt{\log 2} \sqrt{(2RT/Mc^2)} \cdot \nu \text{ cm.}^{-1},$$

where *M* is the molecular weight, *R* the gas constant, *c* the velocity of light, *T* the absolute temperature and ν the wave-number of the spectrum line. Inserting numerical values

$$b = 0.71 \times 10^{-6} \sqrt{(T/M)} \cdot \nu \text{ cm.}^{-1},$$

M, of course, having the value 13 for CH.

Since the half breadth of the line is proportional to the square root of the temperature the method will clearly not be a sensitive or accurate one for temperature

measurement, but by measuring the half breadth of the lines of CH it should at least be possible to determine whether or not there is any marked departure from expected flame temperature and to learn something of the time required for the translational energy of excited radicals to reach equilibrium with the surrounding gas. For obtaining a value for an effective translational temperature it is sufficient to measure the half-width of the line, but to find whether there is any departure from a Maxwell-Boltzmann distribution of translational energy among the molecules it is necessary to compare the intensity contour of the spectrum line with that expected theoretically for pure Doppler broadening.

In order to study the Doppler broadening it is clearly necessary to examine the spectrum line with very large dispersion and to photometer its intensity accurately.

Apart from the question of broadening, the contour of a spectrum line may be modified by self absorption. In the case of CH it is known that flames do not readily show any absorption of these bands, and at the low pressures used in these experiments we may safely ignore self-reversal effects.

Apparatus

For determining the intensity contour of the lines a Fabry-Perot interferometer crossed with a large quartz Littrow-type spectrograph was used.

The interferometer plates, 6 cm. in diameter, were of quartz, and were aluminized by evaporation, the transmission used for most of the work being about 5 %, with a reflexion coefficient of 81 %. The separation between the plates was varied, etalons of 3, 4, 5 and 8 mm. being used, but the 5 mm. one was found most suitable and was used for most of the final measurements. The theoretical resolving power for 81 % reflexion and a 5 mm. separator is 0.07 cm.^{-1} ; in practice the instrumental half-breadth of the lines was around 0.12 cm.^{-1} , although values down to 0.08 were obtained sometimes with short exposure times. Some difficulty was experienced with temperature control, owing to the change of the refractive index of the air between the interferometer plates. The temperature of the room was controlled thermostatically to about 0.5° C and the interferometer itself was placed in a lagged box; in this way it was usually possible to maintain the temperature constant to within 0.1° C for about an hour, which was adequate, but for longer exposures temperature changes were troublesome.

The quartz spectrograph has been described (Gaydon & Wolfhard 1948*a*). It gave excellent definition and had a dispersion of $9 \text{ \AA/mm.}$ in the violet, with an aperture of about $F/25$. Since it was necessary to resolve the rotational fine structure of the CH bands the slit width had to be restricted to 0.2 mm. , which was sufficient to enable the interferometer fringes to be photometered.

The experimental lay-out is shown in figure 1. The technique of running the low-pressure flames has been described (Gaydon & Wolfhard 1948*b*). Light from the flame F is brought roughly parallel by the lens L_1 , passes through the Fabry-Perot interferometer I , and is brought to focus by the lens L_2 on to the slit S of the spectrograph. L_1 was a plano-convex lens of focal length 37.5 cm. and was adjusted to give an image of the flame on the slit when L_2 was in position. L_2 was a good quartz lens (taken from a spectrograph) of focal length 60 cm. For focusing and for determining

the instrumental line width as given by the interferometer a water-cooled mercury arc, *A*, was used; this was placed behind the flame and could be brought into focus on the slit by moving L_1 . The circular interference fringes formed in the interferometer were brought to a sharp focus on the slit by moving the lens L_2 ; since this lens was not achromatic it was necessary to take focus plates for near the wave-lengths required using the water-cooled mercury arc. The interferometer plates were adjusted for parallelism in the usual way, and the interferometer was set at a slight angle so as to throw the centre of the fringe system just above or below the slit which was then crossed by about ten horizontal fringes.

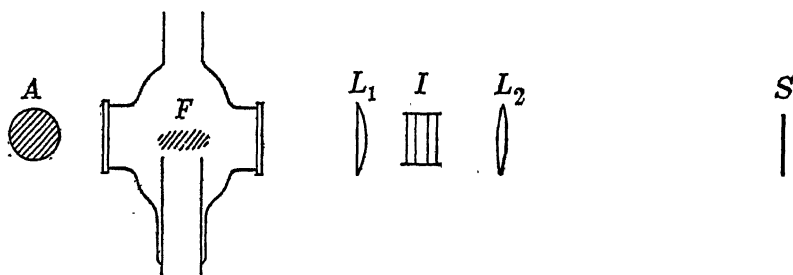


FIGURE 1. The experimental arrangement.

The photometer

A non-recording microphotometer of the usual type with photocell and galvanometer was used. Each photographic plate was calibrated by taking a spectrum (without the interferometer in) of a mercury discharge tube with a stepped sector over the slit; for this a quartz lens was placed over the slit of the spectrograph and the position of the discharge tube was adjusted to give an image in the spectrograph lens, this system ensuring even illumination of the slit; a check exposure with the sector stationary was always taken to ensure that uniform illumination of the slit was in fact being obtained. From the intensity marks made in this way with the sector, the relation between the optical density of the photographic plate and the intensity of the light was derived and this enabled the actual intensity contour of the interference pattern of the spectrum lines from the flame to be obtained. There was usually a faint continuous background in the spectrum of the flame, and the intensity of this was measured in a region close to the spectrum line being studied and this intensity was subtracted to get the true intensity contour of the line.

For all measurements Ilford Zenith plates were used. They had the disadvantage, common to all fast plates, of being rather grainy, but the light-gathering power of the combined interferometer and spectrograph was so small that fast plates were necessary.

The method of computing

The observed intensity contour of a spectrum line results from a combination of the true contour of the line and that given by the interferometer fringe system. In order to get the instrumental contour the water-cooled mercury arc was used; the Doppler broadening of the mercury lines was very small compared with that of the CH lines in the flame because of the water-cooling and the high atomic weight of

mercury; apart from main resonance lines, pressure broadening also appeared small in this arc; most of the mercury lines showed a definite hyperfine structure, but the central component due to the isotopes of even atomic weight was much stronger than other components and no difficulty was found in deriving the instrumental line contour from measurement of suitable mercury lines.

In determining the instrumental contour in this way, it was desirable to do so under exactly the same conditions as those under which the spectrum of the flame was examined, so as to avoid errors due to change in the temperature or adjustment of the interferometer or atmospheric pressure during the exposure. We therefore exposed the mercury arc at the same time as the flame; the mercury lines were thus superposed on the flame spectrum, fortunately falling between the CH lines. In practice, since the arc was brighter than the flame, we made several short exposures of the mercury spectrum at regular intervals throughout the main exposure on the flame.

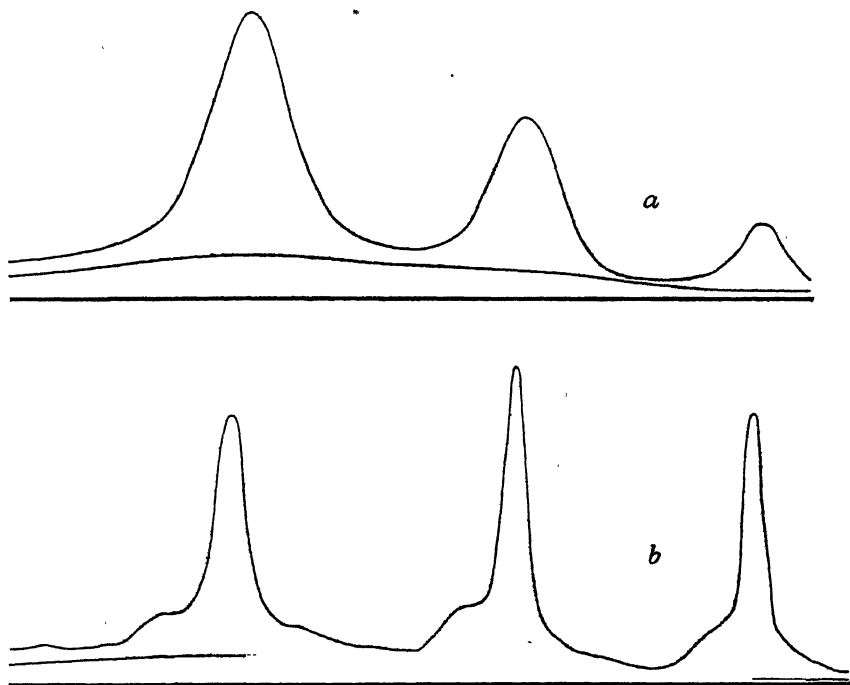


FIGURE 2. (a) Intensity contour of three fringes of the $Q_2(\frac{1}{2})$ line of the $3900\text{\AA}$ band of CH in a stoichiometric oxy-acetylene flame at 9.5 mm. pressure. The lower curve is the continuous background from the flame. (b) The intensity contour of the Hg line at $4077.8\text{\AA}$ under identical conditions; the lower curve again represents the background from the flame. In this plate the fringes are rather more widely spaced than was usually the case, so as to give maximum detail for study of the fringe contour.

The next step was to derive the true half-breadth b , of the spectrum line from the measured half-breadth m , and the instrumental fringe breadth i . For the summation of two curves of the type for collision broadening it is known that $m = b + i$. For the summation of two contours of Doppler-broadening type, however, we have the relation $m^2 = b^2 + i^2$. For the summation of a Doppler contour and the instrumental

contour we are unaware of any exact law, and the resultant seems to be between that given by the square and by the linear summation. Owing to varying effects of adjustment, focus, and temperature change, the instrumental contour changes slightly from plate to plate. It was therefore necessary to make a detailed calculation for each plate.

An ideal Doppler contour for an assumed temperature was plotted using the expression $I = \exp[-\{2\sqrt{(\log_e 2).x/b\}^2}]$, where x is the difference in wave-number from the centre of the line. A graphical integration was then made with this curve and the observed instrumental contour, as derived from the mercury line. The resulting curve was compared with the observed intensity contour. The method was rather laborious, but with practice it was possible to get quite a fair fit from the first approximation, and a good fit with a second approximation.

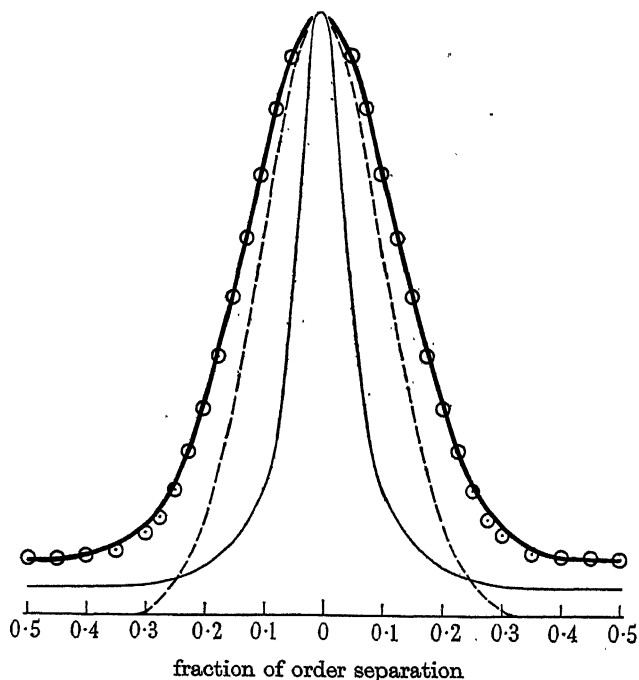


FIGURE 3. — the instrumental contour as deduced from the contour of the Hg line shown in figure 2b. ---- an ideal Doppler curve for CH at 3700° K. — the graphical integration of the two curves. ○ points derived from the observed intensity contour of the strongest fringe in figure 2a. This fringe lies slightly below the centre of the reaction zone of the flame, which accounts for the temperature being a little lower than the value of 4000° K usually obtained. Plate 1 shows a reproduction of the spectrogram from which the curves of figures 2 and 3 were derived.

To obtain accurate intensity contours it was found necessary to make rather heavy exposures so that both the continuous background from the flame and the instrumental background between the interference fringes should be easily measurable; this was only the case when these backgrounds were well above the threshold of sensitivity of the photographic plate.

Figure 2 *a* and *b* and figure 3 show respectively a typical group of observed intensity contours for a CH line, the observed contour of the mercury line at 4077.6 Å for deriving the instrumental contour, and in figure 3 three curves, one the deduced instrumental contour, the second the assumed line contour for a selected value for the temperature, and the third the result of graphically integrating these two; points derived from the experimentally obtained curve are plotted along the calculated curve for comparison.

The spectrum lines used

The CH band at 3900 Å corresponds to a $^2\Sigma \rightarrow ^2\Pi$ transition, and the band has *P*, *Q* and *R* branches, each branch consisting of double lines due to the spin splitting in the $^2\Pi$ state. This spin splitting decreases with increasing rotational energy and the lines become close together and cannot be resolved with the relatively wide slit which must be used for photometering along the spectrum line. In consequence, there was only one line, the weak $Q_2(\frac{1}{2})$ line, in the 3900 Å band which was suitable for detailed study. The weakness of this line and the need to bring up the background between the fringes above the threshold of plate sensitivity meant that heavy exposures were required. The reduction of light by the interferometer and spectrograph therefore limited observations to very bright flames, and in fact satisfactory results were only obtained with oxy-acetylene flames. Exposure times were usually around $\frac{1}{2}$ hr., although some exposures up to 2 hr. were made; with the longer exposures temperature changes in the interferometer were a limiting factor.

For the CH band at 4315 Å, due to a $^2\Delta \rightarrow ^2\Pi$ transition the rotational structure was even more complex the *P*, *Q* and *R* branches each having four close components due to spin and Λ -doubling. Near the band head the structure was impossibly complex, but near the tail of the band where only the *R* branches remained strong the Λ -doubling was well resolved, while the spin doubling was quite small. For a few lines of the R_{cd} branch the components R_{1cd} and R_{2cd} were separated by about 1 cm^{-1} and with suitable separation of the interferometer plates the fringes of the two lines were roughly equally separated. In this way the combined fringe system could be microphotometered and calculations made as for the line of the 3900 band but with the graphical integration complicated by the need to take into account the presence of two lines of equal strength. The resolving power was also reduced by the need to separate the fringes and hence to use a smaller separator. The majority of measurements were therefore made with the line of the 3900 Å band, but some were made as well with the 4315 Å band.

TRANSLATIONAL TEMPERATURES; RESULTS

Oxy-acetylene flame at atmospheric pressure

For an ordinary welding flame the continuous background was very strong when the spectrum was photographed using a wide slit, and accurate results could not be obtained. Using a 3 mm. etalon, to increase the sharpness of the fringes, an approximate value of 2600° K was obtained, using the $Q_2(\frac{1}{2})$ line of the 3900 Å band. The exact mixture strength in this flame was not known, but it was probably not far

from stoichiometric and the theoretical flame temperature would be rather higher than the observed translational temperature in the reaction zone as obtained in this way. The value does, however, serve to show that even at atmospheric pressure the pressure broadening of the CH lines is comparatively negligible and that at this pressure the average translational temperatures of the excited molecules does reach approximate equilibrium with the gas temperature.

Flames at low pressure; the centre of the reaction zone

In the flames at low pressure the reaction zone varies from a few mm. up to about 3 cm. thick, according to the pressure, and it was possible to determine the variation of translational energy through the zone. For the moment we shall concentrate on measurements for part of the reaction zone from which the emission of the CH bands was strongest; this lies rather above the centre of the luminous reaction zone as observed visually. Table 1 contains a collection of results from individual plates for stoichiometric oxy-acetylene mixtures using the $Q_2(\frac{1}{2})$ line of the 3900 Å band.

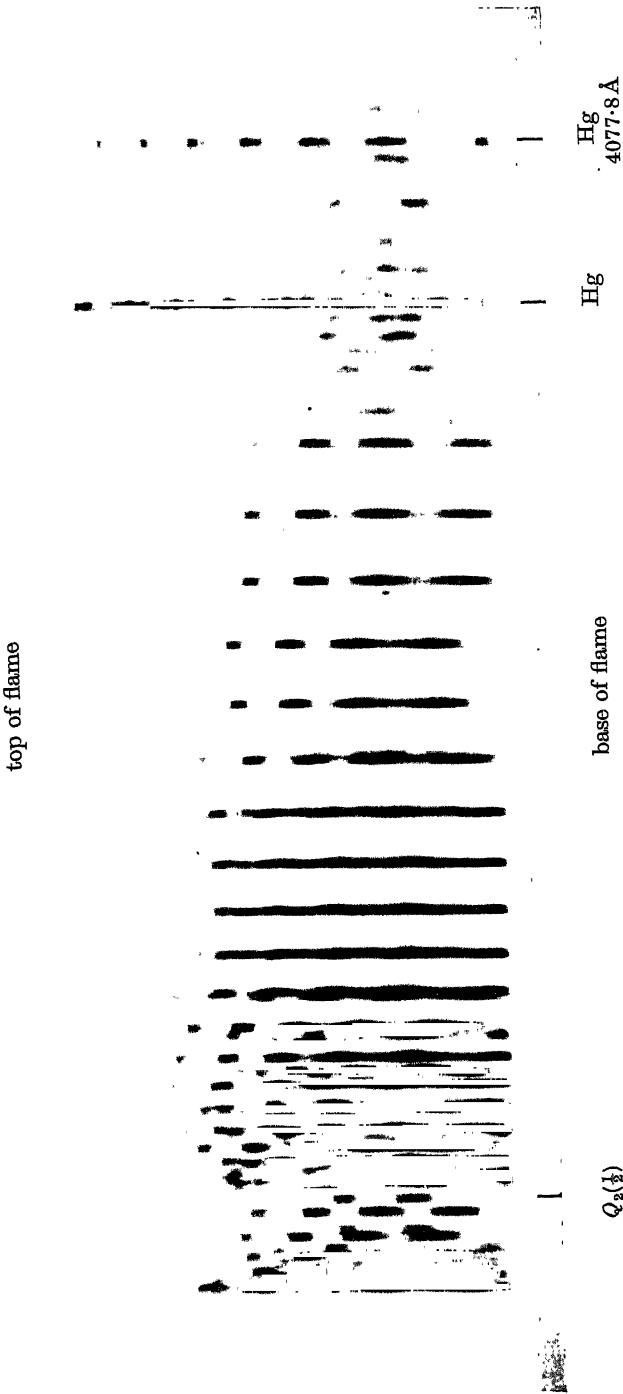
TABLE 1. TRANSLATIONAL TEMPERATURES FOR STOICHIOMETRIC MIXTURES AT VARIOUS PRESSURES

diameter of flame burner (mm.)	gas pressure (mm.)	etalon distance (mm.)	measured half-breadth of CH (cm. ⁻¹)	half-breadth of Hg line (cm. ⁻¹)	temperature (°K)
13	14	3	0.438	0.174	4000
13	14	4	0.407	0.123	4050
13	14	5	0.362	0.117	4150
21	9.5	4	0.347	0.133	3700
21	9.5	5	0.373	0.127	3900
34	5	3	0.42	0.17	4150
34	5	5	0.35	0.112	4000
34	4.2	5	0.374	0.16	4000
54	2.2	5	0.40	0.162	4000
54	2.0	3	0.43	0.185	4500
54	1.9	5	0.37	0.12	4000

Over this pressure range from 2 to 14 mm. there does not appear to be any systematic change of temperature with pressure. Values obtained with the 5 mm. etalon are rather more consistent than with the smaller separators. The mean value appears to be about 4000° K. The scatter of the results suggests a probable error of less than 200° K, but in view of the possibility of more systematic errors it is probably safer to assume a temperature of $4000 \pm 400^\circ \text{K}$ for the centre of the reaction zone.

For mixtures other than stoichiometric the temperatures came only very slightly lower. Thus a rich mixture, containing about double the required amount of acetylene gave 3800° K. A weak mixture containing about 50 % too much oxygen gave 3900° K.

Measurements on stoichiometric mixtures using a line of the 4315 Å band gave less consistent results, values of 3700, 4500, 4000 and 4150° K being obtained from separate plates, but the mean value is again not far from 4000° K.



Spectrum of acetylene-oxygen flame at 9.5 mm. pressure showing the CH band at 3900 Å.
The spectrograph is crossed with a Fabry-Perot interferometer with plates 4 mm. apart.

The observed intensity contour for the line of the 3900 Å band was in satisfactory agreement with the contour calculated from the instrumental contour combined with an ideal Doppler broadening. The degree of agreement varied somewhat from plate to plate, largely because of the experimental error in allowing for the background intensity. In the curves reproduced in figure 3 it will be seen that the circles, derived from the observed line contour, follow the calculated contour quite well. The experimental error is such that the comparison between the observed and calculated curves cannot give high accuracy, but it does seem that there is no marked indication that the distribution of velocities in the excited CH radicals differs appreciably from a Maxwell-Boltzmann distribution at an effective temperature of about 4000° K.

Variation through the reaction zone

The translational temperature as deduced from the Doppler broadening was appreciably higher at the top of the reaction zone and lower at the base of the zone. The results for a typical plate are listed in table 2, for a stoichiometric flame at 2.2 mm. pressure. The actual results varied somewhat from plate to plate owing to the experimental error, but it was clear that at the very base of the reaction zone the effective translational temperature was relatively low, values between 2500 and 3200° K being obtained, while near the top of the reaction zone it was a good deal higher, some plates giving temperatures in excess of 5000° K at the top.

TABLE 2. VARIATION OF TRANSLATIONAL TEMPERATURE
THROUGH THE REACTION ZONE

fringe no.		temperature (° K)
1	top	4400
2		4550
3		4100
4	centre	4050
5		3500
6		3450
7		3200
8	base	3050

MEASUREMENT OF ROTATIONAL TEMPERATURE

The method of calculating the effective rotational temperature from the rotational fine structure of a band is well known. If I is the intensity of a single spectrum line of wave-number ν , then

$$I = CP\nu^4 \exp[-E_r/kT],$$

where C is a constant depending on the total intensity, P is the transition probability for the particular line, E_r is the rotational energy, k is the Boltzmann constant and T is the absolute temperature. From this we have

$$\log_e I - \log_e (P\nu^4) = E_r/kT + \text{constant},$$

hence if we plot $\log I - \log(P\nu^4)$ against E_r/k the slope will give the temperature. I is obtained experimentally from photometry of the spectrum; the transition probability may be derived theoretically.

The quartz Littrow-type spectrograph was again used, and the intensity marks to calibrate the photographic plate were again made with the rotating stepped sector. Over the region of the CH bands it was found, by measurements with a tungsten strip lamp at known temperature using Wien's law, that the effect of change of sensitivity of the plate with wave-length and the change of aperture of the spectrograph was quite negligible for the comparatively short regions in which the lines of the rotational structure were being photometered.

The 3900 Å band of CH

This is due to a $^2\Sigma \rightarrow ^2\Pi$ transition, and the lines of the *P*, *Q* and *R* branches show a small but resolved spin-splitting. In order to resolve this it was necessary to use a fine slit and to be careful to obtain the maximum resolving power from the spectrograph. Temperature changes were troublesome, but could be overcome for exposures up to an hour by using thermostatic control. The use of the fine slit reduced the accuracy of the measurements a little, because of the graininess of the plates, but satisfactory results were obtained. Owing to the predissociation in this band setting in around $K' = 15$ only lines of lower K' could be used.

For the lines used, all with K' less than 14, it was sufficient to take the rotational energy E_r as being equal to $B'K'(K' + 1)$ where B' is the rotational energy constant for the excited $^2\Sigma$ state. Values for the transition probability P were calculated from the formulae given by Earls (1935); there seems to be some inaccuracy in these formulae at very low quantum numbers, the *R* lines being too strong and the *P* lines too weak, so that it was necessary to neglect points with K' less than 4.

The 4315 Å band of CH

This band is due to a $^2\Delta \rightarrow ^2\Pi$ transition. Near its head the structure was much too close to resolve with the dispersion used, but away from the head the lines of the *R* branch were satisfactory. For these the Λ -doubling was easily resolved, but the spin-splitting was quite small (about 1 cm^{-1} in one component, and much less in the other) and so a comparatively wide slit was used and the two spin components were measured together. The Λ -doubling was then not readily resolved before the *R* (12) lines, and points above *R* (22) were disregarded because of the onset of predissociation at *R* (24).

For this band the rotational energy was derived from the expression

$$E_r = B'K'(K' + 1) - D'K'^2(K' + 1)^2,$$

where B' and D' are the rotational energy constants for the $^2\Delta$ state.

Accurate intensity factors for $^2\Delta \rightarrow ^2\Pi$ transitions intermediate between Hund's coupling cases *a* and *b* have not been derived, but both electronic states of CH are very close to case *b* and it was sufficient to use the formulae given by Mulliken, as summarized by Jevons (1932).

The 3143 Å band of CH

This is much weaker than the other two bands and is usually at least partially masked by the OH band. However, we were able to make measurements for a few flames. The transition is $^2\Sigma \rightarrow ^2\Pi$ like the 3900 Å band. The spin-splitting was rather

small, and so a rather wide slit was used and the two spin components were measured together. Data for the rotational analysis were taken from T. Heimer (1932) and a number of lines of the *P* and *R* branches which were not overlapped by OH lines were used.

VALUES FOR EFFECTIVE ROTATIONAL TEMPERATURES

Results for the 3900 Å band

Owing to the need to work with the large quartz spectrograph and to use some of the weaker lines of the *R* and *P* branches, it was only possible to apply this method of measuring the rotational temperatures to relatively bright flames. All the mixtures used were stoichiometric. In all cases a satisfactory straight line was obtained, apart from the discrepancy between the *P* and *R* lines at very low rotational energies already mentioned and believed to be due to failure of the intensity formulae at very low values of *K*. The fact that the graph was fairly straight indicated that the distribution of rotational energy among the excited CH radicals was similar to that for a Maxwell-Boltzmann distribution. Figure 4 shows a typical graph of

$$\{\log I - \log (P\nu^4)\}$$

against E_r , and table 3 summarizes the results, all measurements being for the part of the reaction zone from which the CH bands were emitted most strongly.

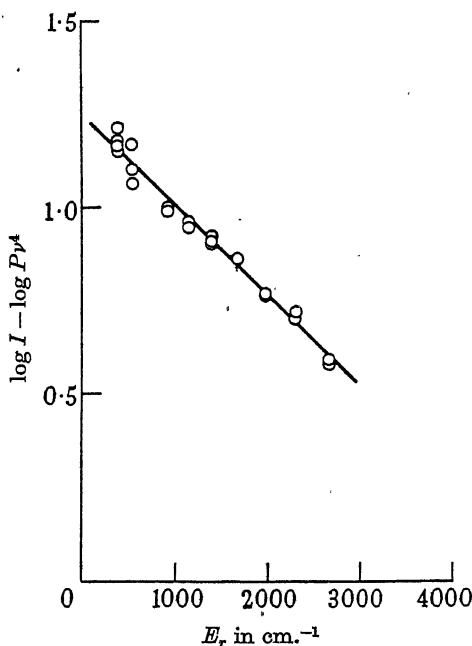


FIGURE 4. Results for the 3900 Å band of CH from the flame $C_2H_2 + 2\frac{1}{2}O_2$ at 6 mm. pressure; temperature = 2600° K.

It will be seen that the values all come near the theoretical flame temperatures. The probable error is about 200° K. There was, however, no marked change of effec-

tive rotational temperature through the reaction zone, as with the translational temperature. For the oxy-acetylene flame at 1.8 mm. there was some indication of a slight rise of temperature passing upwards through the reaction zone, but it was quite small, being about 2350° at the base, 2460° at the brightest part, and 2580° K at the top.

TABLE 3. EFFECTIVE ROTATIONAL TEMPERATURES FOR CH 3900 Å

flame	pressure	temperature (° K)
oxy-acetylene welding flame	1 atm.	3200
oxy-acetylene low-pressure flame	6 mm.	2520
	1.8 mm.	2460
	18 mm.	2670
oxy-methane	8 mm.	2800

Results for the 4315 Å band

This band again gave fairly good straight lines for the plot of $\{\log I - \log (P\nu^4)\}$ against E_r as shown by the examples in figure 5. These are for individual plates and the values differ slightly from the averaged values summarized in table 4; these are again for the brightest part of the reaction zone. Results were again consistent to about 200° K or perhaps a little better in some cases.

These results show that the effective rotational temperatures for the CH radicals in the excited $^2\Delta$ state do depend on the theoretically expected flame temperatures,

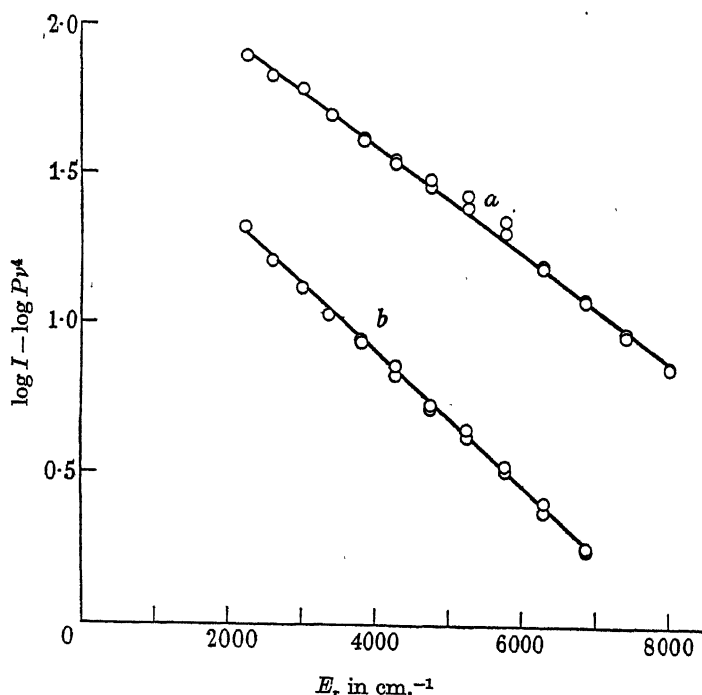


FIGURE 5. Results for the 4315 Å band of CH. *a*, welding flame; $T = 3500^\circ \text{K}$;
b, $\text{C}_2\text{H}_2 + 2\frac{1}{2}\text{O}_2$ at 1.7 mm. pressure; $T = 2690^\circ \text{K}$.

but tend to come rather on the high side, being in some flames up to 10 % too high. The result for methyl alcohol seems in line with the results for acetylene. This methyl alcohol flame was examined because it is one of the few flames for which the CH bands are reasonably strong but which does not show strong C_2 bands. For the 4315 Å band there was little variation in the temperature from the base to the top of the flame.

TABLE 4. EFFECTIVE ROTATIONAL TEMPERATURES FOR CH 4315 Å

flame	pressure	temperature (° K)
$C_2H_2 + 2\frac{1}{2}O_2$	1.7 mm.	2750
$C_2H_2 + 1.7O_2$	2.2 mm.	2750
$C_2H_2 + 5O_2$	2.2 mm.	2300
$C_2H_2 + 12$ air	12 mm.	2420
$C_2H_2 + 20$ air	21 mm.	2050
CH_3OH & O_2 approx. stoichiometric	8 mm.	2400
oxy-acetylene welding flame	1 atm.	3600

Results for the 3143 Å band

This band was only obtained in sufficient strength and not overpowered by the overlapping OH in flames of rather fuel-rich mixtures of oxygen and acetylene. Fairly good straight lines were obtained for lines of both the *P* and *R* branches, but the *P* lines were relatively too strong compared with the *R* lines.

For the welding flame at 1 atm. pressure a value of 3700° K was obtained; this is appreciably higher than the theoretical maximum temperature for the rich mixture used. For the low-pressure flame, at 11 mm., the very high value of 5070° K was obtained. It thus seems that this band behaves differently from the other two. The band is emitted from the same part of the reaction zone of the low-pressure flame, the upper part, as the other bands of CH. There is, however, some indication that it is favoured, relative to the other bands, by richer mixtures. Also its occurrence is restricted to very hot flames. The overlapping by the OH makes accurate observations of its relative intensity difficult.

DISCUSSION OF RESULTS

The experimental results may be summarized as follows. The effective translational temperatures for both the 3900 and 4315 Å bands in the low-pressure oxy-acetylene flames come much too high, being around 4000° K against maximum theoretical temperatures of 2500 to 2800° K. The translational temperature is appreciably lower at the base of the reaction zone and higher still near the top. At atmospheric pressure the effective translational temperature is near the theoretical maximum flame temperature. The effective rotational temperature for the first excited $^2\Sigma$ state of CH (3900 Å band) is near the theoretical maximum flame temperature, both at atmospheric and at low pressure; there are indications of a slight rise from the base to the top of the reaction zone. For the excited $^2\Delta$ state (4315 Å band) the effective rotational temperature is again closely related to the flame temperature, but tends

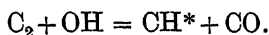
to come up to 10 % high both at atmospheric and at low pressure, and there is little indication of any variation from the base to the top of the reaction zone. The second excited $^2\Sigma$ state, the upper state of the weak 3143 Å band, gives quite a different result, with a rotational temperature of 3700° K at 1 atm. and over 5000° K at low pressure.

We thus have the rather surprising result that for the main bands at 3900 and 4315 Å the rotational energy of the excited CH molecules is closely related to the gas temperature, while the translational energy in the low-pressure flames is considerably in excess of what must be the average translational temperature of the gas molecules. For the case of OH it was shown (Gaydon & Wolfhard 1948*a*) that the rotational temperature was abnormally high and that this was due to the OH radicals being formed in the excited electronic state by chemical reaction and to the number of collisions during the lifetime of the excited radical being too few to bring the rotational temperature into equilibrium with that of the surrounding molecules. For CH we should expect a shorter radiative life than for OH (it having been shown that OH has a long radiative life of the order 4×10^{-6} sec.). There seem two possible interpretations, first to assume that the CH radiation is not due to chemiluminescence, or secondly to assume that the chemiluminescence has little energy to spare for high rotational excitation of the newly formed CH radical. The high translational temperature in the low-pressure flames clearly shows that the excitation of the CH radicals does not result from normal thermal causes and also shows that at pressures below 14 mm. the number of collisions occurring during the radiative life of the excited radicals is too few for excess translational energy to be lost to surrounding molecules. We should expect that almost every collision would lead to a considerable loss of excess translational energy and that something near equilibrium would be attained in a very few collisions. At 14 mm. pressure and around 2500° K one might expect around 5×10^6 collisions per sec. Since in this range the translational temperature does not vary appreciably with pressure we may conclude that the radiative life of the excited CH is less than 2×10^{-7} sec. On the other hand at atmospheric pressure, where kinetic theory data would indicate around 3×10^9 collisions per sec., translational energy equilibrium appears to be established, and hence the radiative life must be appreciably longer than 3×10^{-10} sec. These values would be in accord with a normal radiative lifetime for the CH radicals around 10^{-8} sec.

If we consider the first assumption, that the excitation of the CH is not due to true chemiluminescence, then since it is clearly not normal thermal excitation we must conclude that the radicals receive their electronic excitation from other active species in the flame, but that in this way there is little conversion to rotational energy, so that the rotational temperature of the CH remains related to the rotational temperature of the surrounding molecules, although the CH gains both electronic and translational energy. If the angular momentum of the system has to be conserved during the collision then it seems reasonable that this may be the case.

If we examine the second possible assumption, that the excitation is due to chemiluminescence, but that it does not result in giving the new CH radicals an appreciable excess of rotational energy, we require a fuller knowledge of the actual chemiluminescent process and of the steric factors to be able to decide whether or not this is likely.

The simplest reaction for the formation of excited CH is



This reaction has been supported in the past by the work of Pannetier & Gaydon (1947) on the production of CH in moist cyanogen flames. There is some doubt about the energies of dissociation of C_2 and CO, but this reaction is certainly very exothermic and we should have expected that it would lead to a considerable excess of rotational energy in the newly formed CH radical. In the low-pressure flames the CH bands usually occur high in the flame while the C_2 bands occur lower. From the above reaction the strength of the CH emission should depend on the concentrations of C_2 and OH; while we might expect the maximum of intensity of the CH to lie higher than that for C_2 , it seems difficult to account for the strength of the CH at the very top of the flame where the C_2 bands seem quite absent.

Apart from the above reaction, it is very difficult to postulate a reasonable reaction which would be sufficiently exothermic to form electronically excited CH. Reactions of C_2 with HO_2 , or of atomic carbon with some other molecule give enough energy but seem improbable.

There is some indication that the rotational temperature obtained from the 4315 Å band is a little (200° K.) higher than from the 3900 Å band. This may be because there is a slight departure from a Maxwell-Boltzmann distribution of rotational energy and the 4315 Å band uses points at relatively high energy, while the curve for the 3900 Å band is limited to points from low rotational energy levels, so that the slight difference may not be significant. If it were, however, shown to be genuine then it might support the view that the excitation is a real chemiluminescence, although the energy difference between the upper levels of the two bands corresponds to much more than 200° K.

There does not, at present, seem to be any conclusive way of deciding between the two possible modes of excitation we have discussed, but on present experimental evidence we favour the view that the excitation of the CH radicals results from collision of unexcited CH radicals with some other excited species in the flame and not from either true chemiluminescence or normal thermal excitation. Our reason for this is the difficulty of finding a satisfactory mechanism for a chemiluminescence, and the knowledge that in other cases there appears to be an unusually high excitation temperature in flames; for example we have found (Gaydon & Wolfhard 1948*a*) that iron atoms show a very high temperature in the reaction zones of acetylene flames. We hope that further measurements on translational and excitation temperatures of other molecules and atoms may settle this problem.

The excitation of the weak 3143 Å band of CH appears to be of a different character from that of the two main bands. The high rotational temperature, increasing to very high values at low pressure, resembles the results for OH, and we are most probably dealing with true chemiluminescence. This band, however, is relatively weak and requires very hot flames for its occurrence. This suggests that the reaction responsible for the chemiluminescence requires a considerable activation energy.

In conclusion we wish to thank Sir Alfred Egerton, F.R.S., for his interest and his suggestion to use Doppler widths to measure translational temperatures, and one

of us (A.G.G.) is grateful to the Royal Society for the award of a Warren Research Fellowship.

REFERENCES

- Bundy, F. P. & Strong, H. M. 1948 *Third symposium on combustion*. Wisconsin.
 Earls, L. T. 1935 *Phys. Rev.* **48**, 423.
 Gaydon, A. G. & Wolfhard, H. G. 1948a *Proc. Roy. Soc. A*, **194**, 169.
 Gaydon, A. G. & Wolfhard, H. G. 1948b *Disc. Faraday Soc.* **2**, 161.
 Gaydon, A. G. & Wolfhard, H. G. 1948c *Colloquium on combustion*. Paris.
 Heimer, T. 1932 *Z. Phys.* **78**, 771.
 Jevons, W. 1932 *Band spectra of diatomic molecules*. Physical Soc.
 Klaukens, H. & Wolfhard, H. G. 1948 *Proc. Roy. Soc. A*, **193**, 512.
 Pannetier, G. & Gaydon, A. G. 1947 *C.R. Acad. Sci., Paris*, **225**, 1300.
 Tolansky, S. 1947 *High resolution spectroscopy*. Methuen.
 Wolfhard, H. G. 1939 *Z. Phys.* **112**, 107.

Distribution of solute atoms round a slow dislocation

BY A. H. COTTRELL AND M. A. JASWON

Metallurgy Department, University of Birmingham

(Communicated by N. F. Mott, F.R.S.—Received 21 March 1949)

The factors determining the average velocities of solute atoms attracted to a dislocation are discussed, and an equation is set up for the concentration of solute round the dislocation. This leads to a symmetrical distribution of a Maxwell-Boltzmann type round a stationary dislocation. In the case of a slowly moving dislocation the distribution is given as a series solution in Mathieu functions, and the non-symmetry of this distribution causes a force on the dislocation, opposing its motion. This force is calculated and shown to increase linearly with the speed of the dislocation, at low speeds, and a critical range exists above which the motion is unstable and the dislocation accelerates. The characteristics of the plastic flow below this critical range are compared with those of micro-creep in tin crystals and are shown to be similar. Quantitative agreement can be obtained by assuming plausible values for the density of dislocations and the rate of diffusion of solute atoms in tin, but the need for further experiments on micro-creep is emphasized.

1. INTRODUCTION

There are grounds for believing that solute atoms which strain a crystal lattice migrate to dislocations and there form equilibrium 'atmospheres' in which large atoms are collected in the regions expanded by the dislocations and small atoms in the compressed regions (Cottrell 1948). An atom in substitutional solution which can relieve hydrostatic stresses is considered to interact with a positive edge dislocation with a potential given by

$$V = \frac{4}{3} G \epsilon a^3 \lambda \frac{1 + \nu \sin \theta}{1 - \nu} \frac{\theta}{r}, \quad (1)$$

where r and θ are its co-ordinates relative to the dislocation, G and ν are the rigidity modulus and Poisson's ratio, respectively, λ is the slip distance in the dislocation,

and r_a and $r_a(1+\epsilon)$ are the respective atomic radii of solvent and solute. The equilibrium concentration of this species of solute in the neighbourhood of the point r, θ is then

$$c(r, \theta) = c_0 \exp\{-V(r, \theta)/kT\}, \quad (2)$$

where c_0 is the average atomic concentration, provided that the dislocation is at rest and that the atmosphere is not so concentrated that the local dilatation caused by solute atoms is comparable with that caused by the dislocation.

The existence of such atmospheres makes it possible to divide moving dislocations into two classes, slow and fast, the slow ones being accompanied by imperfectly formed atmospheres and the fast ones moving too fast to possess atmospheres. The speed of a slow dislocation is limited by the rate of migration of the solute atoms, whereas a fast dislocation probably moves at a speed approaching that of sound (Mott & Nabarro 1948; Frank 1948); it seems unlikely that stable states of motion at intermediate speeds can exist. In a preliminary analysis (Cottrell 1948) it was shown that slow dislocations could move fast enough to account for the rates of creep observed in certain slow creep tests and, in particular, in the micro-creep of tin crystals observed by Chalmers (1936). It was also suggested that the non-symmetrical distribution of solute atoms round a moving dislocation would cause a perturbation force to act on the dislocation in opposition to the force due to the applied stress; this force can be regarded as a viscous drag which increases as the speed of the dislocation increases, for small speeds, thereby allowing a steady state of motion to be reached under a steady applied stress. The concept of slow dislocations has also been discussed by Nabarro (1948) in connexion with the mechanical properties of iron at elevated temperatures.

The purpose of this paper is to develop a quantitative theory of the way that solute atoms distribute themselves round a slow dislocation, and to examine the characteristics of the viscous drag caused by this distribution. The theory is confined to the case of a positive-edge dislocation interacting with solute atoms by means of the hydrostatic effect represented by equation (1), and it is assumed that the equilibrium atmosphere is Maxwellian (equation (2)), i.e. that the solution is dilute and the interaction is weak. This means that the theory cannot be applied quantitatively to atmospheres of carbon and nitrogen in iron, where the strains round solute atoms are not even approximately spherically symmetrical (Nabarro 1948) and the atmosphere is not Maxwellian (Cottrell & Bilby 1949). However, it is believed that the assumptions made allow the deductions of the theory to be compared with the behaviour observed in micro-creep.

2. THE EQUATION OF CONCENTRATION

A positive-edge dislocation is assumed to be moving with speed v in the direction of slip; the direction of v defines the positive x -axis in a right-handed system and the dislocation line coincides with the z -axis (figure 1). The differential equation satisfied by the distribution of the solute atoms round the dislocation is obtained by considering that a group of solute atoms in a small region of the crystal acquires a velocity relative to the dislocation in three ways:

(i) A drift velocity due to the field of force $-\nabla V$; this is given by Einstein's formula for Brownian motion as

$$\mathbf{v}_1 = D\mathbf{F}/kT = -D\nabla V/kT, \quad (3)$$

where D is the diffusion coefficient.

(ii) A diffusion velocity caused by the concentration gradients set up by the segregation of the solute atoms. From Fick's law this is

$$\mathbf{v}_2 = -D\nabla c/c. \quad (4)$$

(iii) A constant velocity $-v$ along the x -axis, due to the motion of the dislocation; it is convenient to think of the dislocation being at rest and the crystal flowing past it with this velocity.

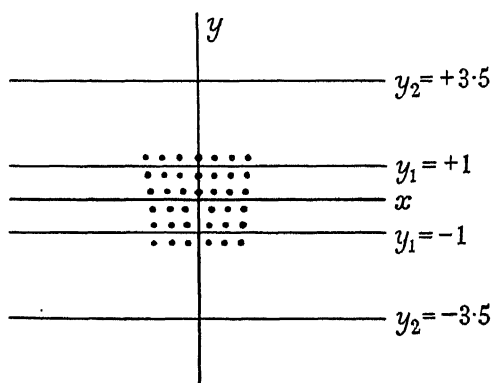


FIGURE 1

The total relative velocity of the solute atoms is therefore

$$\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 - v\mathbf{i}, \quad (5)$$

where $\mathbf{i}$ is the unit vector along the x -axis.

In a steady state of motion the concentration in any small region at a fixed position relative to the dislocation must not change with time, so that the equation of continuity to be satisfied is $\nabla(c\mathbf{v}) = 0$. On substituting for $\mathbf{v}$ and writing

$$U = V + (kT/D)vx,$$

this becomes

$$\nabla(c\nabla U + kT\nabla c) = 0, \quad (6)$$

which is the required equation of concentration. This equation is a second-order partial differential equation of the elliptic type for c as the unknown with the boundary condition that $c = c_0$ at infinity. If $v = 0$, then $U = V$, and the equation determines the concentration round a dislocation at rest. In this case an additional condition is that there should be no resultant velocity anywhere, so that only the particular solution $c\nabla V + kT\nabla c = 0$ is required; this is readily integrated to give equation (2).

When the dislocation is moving the solution has to satisfy three conditions. It must reduce (i) to $c = c_0 \exp(-V/kT)$ when $v \rightarrow 0$, (ii) to $c = c_0$ everywhere when

$V \rightarrow 0$ everywhere, and (iii) to $c = c_0$ at infinity. By analogy with equation (2), one might suggest the pseudo-equilibrium solution $c \nabla U + kT \nabla c = 0$, which leads to $c = c_0 \exp(-U/kT)$, for the moving dislocation, but it fails to satisfy conditions (ii) and (iii). It corresponds to the case where the solute atoms have no resultant velocity relative to the dislocation, which obviously only holds when the dislocation is at rest.

To find a more general solution of equation (6) it is necessary to consider the explicit form of V . Equation (1) gives $V = A \sin \theta / r$, where $\theta = \tan^{-1}(y/x)$ and $r = (x^2 + y^2)^{1/2}$, in which case the concentration has a singularity at the origin. This is because equation (1) is based on the elasticity theory of a dislocation in a continuous medium, which is not applicable at the origin. It is to be expected on physical grounds that the dilatation is zero at the centre of a dislocation, and this is supported by Nabarro's theory of dislocations (1947) which shows that, when account is taken of the discontinuous atomic structure of the medium, the stress field does not become infinite at the origin. However, the simple form $A \sin \theta / r$ can be retained in the present work because the boundary conditions, which together with equation (6) determine the concentration, do not involve the origin, and because the stresses given by the simple theory approximate closely to those given by Nabarro except within a region of about two atoms from the origin. As will be seen later, conditions at the origin are not very important in the present problem; the rate of adjustment of the atmosphere to equilibrium is rapid near the origin, so that the perturbation of concentration is small in this region.

3. SEPARATION INTO PARTS

By setting $c/c_0 = \psi \sqrt{\phi}$, where $\psi(x, y)$ is a function to be determined and

$$\phi = \exp(-U/kT) = \exp(-V/kT) \exp(-vx/D) \quad (7)$$

is the pseudo-equilibrium function, equation (6) can be transformed to

$$\nabla^2 \psi - \frac{\nabla^2 \sqrt{\phi}}{\sqrt{\phi}} \psi = 0, \quad (8)$$

which, in polar co-ordinates, is separable into parts. The boundary conditions for ψ are (i) $\psi = \exp(-V/2kT)$ when $v = 0$, (ii) $\psi = \exp(vx/2D)$ when $V = 0$, and (iii) $\psi = O[\exp(vx/2D)]$ at infinity. Equation (8) could be written out explicitly and separated as it stands, but it is convenient first to introduce the changes $l = A/kT$, $\rho = r \sqrt{(v/Dl)}$, $q = vl/4D$ and $\omega = \theta - \frac{1}{2}\pi$, in which case it becomes

$$\rho \frac{\partial}{\partial \rho} \left(\rho \frac{\partial \psi}{\partial \rho} \right) + \frac{\partial^2 \psi}{\partial \omega^2} - (q\rho^2 + q/\rho^2 - 2q \cos 2\omega) \psi = 0. \quad (9)$$

The factors ρ and q are non-dimensional. It will be shown later, on the basis of the magnitudes of the physical quantities involved, that the perturbation parameter q is a small fraction.

Introducing $\psi = P(\rho)\Omega(\omega)$ in the equation and separating into parts in the usual manner, the angular function Ω is found to satisfy the Mathieu equation

$$\frac{d^2 \Omega}{d\omega^2} + (\alpha + 2q \cos 2\omega) \Omega = 0, \quad (10)$$

and the equation for the radial function P is

$$\rho \frac{d}{d\rho} \left(\rho \frac{dP}{d\rho} \right) - (a + q\rho^2 + q/\rho^2) P = 0, \quad (11)$$

where a is a constant. Since the concentration must be single-valued the angular function must be periodic in 2π and, from the theory of Mathieu functions (Whittaker & Watson 1946), this necessitates that a be a function of q , of the form $a = n^2 + O(q)$, where n is any positive integer or zero.

When q is zero, equation (10) reduces to a harmonic equation for Ω , with $a = n^2$, having the solutions

$$\begin{array}{ccccccc} 1, & \cos \omega, & \cos 2\omega, & \dots, & \cos n\omega, & \dots, \\ & \sin \omega, & \sin 2\omega, & \dots, & \sin n\omega, & \dots \end{array}$$

Following McLachlan (1947), the analogous solutions of equation (10) are

$$\begin{array}{ccccccc} ce_0, & ce_1, & ce_2, & \dots, & ce_n, & \dots, \\ & se_1, & se_2, & \dots, & se_n, & \dots, \end{array}$$

where $ce_n = \cos n\omega + O(q; \omega)$ and $se_n = \sin n\omega + O(q; \omega)$, and $O(q; \omega)$ vanishes when $q = 0$. Since $q = vl/4D$ the angular Mathieu functions reduce to the corresponding harmonics when $v = 0$ or $l = 0$. It should also be noted that the harmonics $\cos n\omega$ and $\sin n\omega$ correspond to the same value of the characteristic number, $a = n^2$, but in the analogous Mathieu functions the harmonic level is split and the characteristic numbers differ to the order q .

The only angular functions which will be needed explicitly in the present problem are ce_1 and se_1 , which are, to the order q ,

$$\begin{array}{l} ce_1 = \cos \omega + \frac{1}{8}q \cos 3\omega, \quad (b_1) = 1 - q, \\ se_1 = \sin \omega + \frac{1}{8}q \sin 3\omega, \quad (a_1) = 1 + q, \end{array} \quad (12)$$

with the corresponding characteristic numbers, to the order q , designated in brackets at the side.

As can be seen by writing $\rho = e^z$, the radial equation is a hyperbolic Mathieu equation and, since the parameter a is the same as in equation (10), each solution of the former equation must correspond to one of the latter. Of the several types of solutions available, the functions Fek_n^* and Gek_n^* will be used. These are multiples of the functions Fek_n and Gek_n , defined in McLachlan (1947), and have the properties:

$$(i) \quad Fek_n^*, \quad Gek_n^* = O(\rho^{-\frac{1}{2}} e^{-\rho\sqrt{q}}) \text{ at infinity (McLachlan 1947, p. 248)}$$

and

$$(ii) \quad Fek_n^* = Gek_n^* = 2I_n(\sqrt{q}/\rho) [1 + O(q \ln q; \rho)],$$

where I_n is the Bessel function of the first kind with imaginary argument, and where, as may be shown (Jaswon 1949), the term $O(q \ln q; \rho)$ contains no terms which are powers of either $\rho\sqrt{q}$ or $\sqrt{q}/\rho$ and vanishes in the limit when q is zero, i.e. when either v or l is zero.

From the fact that $\sqrt{q}/\rho = l/2r$ and from the properties of $O(q \ln q; \rho)$ just noted, it follows that

$$[Fek_n^*]_{v \rightarrow 0} = [Gek_n^*]_{v \rightarrow 0} \rightarrow 2I_n(\sqrt{q}/\rho),$$

and, when $n > 0$,

$$[\text{Fek}_n^*]_{l \rightarrow 0} = [\text{Gek}_n^*]_{l \rightarrow 0} \rightarrow [2I_n(\sqrt{q}/\rho)]_{l \rightarrow 0} = 0,$$

since $I_n(x) \rightarrow 0$ as $x \rightarrow 0$ except when $n = 0$.

4. A SERIES SOLUTION

The simplest general solution of equation (9) which satisfies all the boundary conditions is

$$\psi = \exp(V/2kT) \exp(vx/2D) - \sum_0^\infty A_{2n+1}^c \text{ce}_{2n+1} \text{Fek}_{2n+1}^* + \sum_0^\infty A_{2n+1}^s \text{se}_{2n+1} \text{Gek}_{2n+1}^*, \quad (13)$$

where the constants A_m^c and A_m^s are the constants associated with ce_m and se_m respectively in the Mathieu function expansion of $\exp(V/2kT) \exp(vx/2D)$. This expansion is a particular case of a more general expansion (McLachlan, p. 206), as can be seen by writing

$$\begin{aligned} \exp(V/2kT) \exp(vx/2D) &= \exp[(\sqrt{q}/\rho) \sin \theta + \sqrt{q}\rho \cos \theta] \\ &= \exp(\cosh z \cos \omega + \sinh z \sin \omega), \end{aligned}$$

where $z = \ln \rho$ and $\omega = \theta - \frac{1}{4}\pi$, and leads to the expressions

$$A_m^c = 2 \cos(m\pi/4) [1 + O(q)]$$

and

$$A_m^s = -2 \sin(m\pi/4) [1 + O(q)].$$

The product pairs $\text{ce}_{2n+1} \text{Fek}_{2n+1}^*$ and $\text{se}_{2n+1} \text{Gek}_{2n+1}^*$ provide particular solutions of equation (9). Also, the function $\exp(V/2kT) \exp(vx/2D)$ may be expanded into a series of Mathieu function product pairs, each pair of which is a particular solution. Hence the function ψ is a solution of the equation of concentration.

Considering the boundary conditions, it should be noted that, at infinity,

$$\begin{aligned} \psi &= [\exp(V/2kT) \exp(vx/2D)]_\infty + O[\rho^{-\frac{1}{2}} \exp(-\rho\sqrt{q})] \\ &= O[\exp(vx/2D)] + O[(x/\cos \theta)^{-\frac{1}{2}} (v/Dl)^{-\frac{1}{2}} \exp(-vx/2D \cos \theta)] \\ &= O[\exp(vx/2D)], \end{aligned}$$

as is required by the third boundary condition.

When $l = 0$ the solution becomes

$$\psi = [\exp(V/2kT) \exp(vx/2D)]_{l=0} = \exp(vx/2D),$$

as is required by the second condition.

When $v \rightarrow 0$, then $\text{Fek}_n^* = \text{Gek}_n^* \rightarrow 2I_n(\sqrt{q}/\rho)$, $A_m^c \rightarrow 2 \cos m\pi/4$, $A_m^s \rightarrow -2 \sin m\pi/4$, $\text{ce}_m \rightarrow \cos m\omega$, $\text{se}_m \rightarrow \sin m\omega$ and therefore

$$\begin{aligned} \psi &= [\exp(V/2kT) \exp(vx/2D)]_{v=0} - \sum_0^\infty 2 \cos[(2n+1)\pi/4] \cos[(2n+1)\omega] 2I_{2n+1}(\sqrt{q}/\rho) \\ &\quad - \sum_0^\infty 2 \sin[(2n+1)\pi/4] \sin[(2n+1)\omega] 2I_{2n+1}(\sqrt{q}/\rho) \\ &= \exp(V/2kT) + (-1)^{n+1} 4 \sum_0^\infty I_{2n+1}(\sqrt{q}/\rho) \sin[(2n+1)\theta], \end{aligned}$$

and, upon forming Bessel function series for $\exp(V/2kT)$ and $\exp(-V/2kT)$ (Copson 1935, p. 321), one obtains $\psi_{v=0} = \exp(-V/2kT)$, as is required by the first boundary condition.

The numerical evaluation of ψ is facilitated by observing that, from the symmetry properties of hyperbolic Mathieu functions with respect to an interchange of ρ and $1/\rho$ (Jaswon 1949) and from the properties of trigonometric functions, evaluating ψ at the point ρ, θ is equivalent to evaluating it at the point $1/\rho, -(\theta + \frac{1}{2}\pi)$. Hence computations are only necessary in the region $\rho \geq 1$. The functions Fek_n^* and Gek_n^* decrease rapidly with increasing index, for $\rho \geq 1$, so that in this region the solution is adequately represented by

$$\psi = \exp(V/2kT) \exp(vx/2D) - A_1^* \text{ce}_1 \text{Fek}_1^* + A_1^* \text{se}_1 \text{Gek}_1^*, \quad (14)$$

which is much easier to compute. It should be noted that the solution could be used for all values of q , although in practice the computations become extremely difficult unless q is a small fraction.

Using equation (14), ψ has been computed at the points $\rho = 0.20, 0.25, 0.5, 1, 2, 4, 5$ and 10 along each of a series of radii emanating from the origin, the radii being taken at intervals of 15° . Taking $q = 0.04$ as a convenient value, accurate computations of the relative concentration, c/c_0 , were then made to provide data for estimating the perturbation force. An approximate calculation was also made for a faster dislocation, $q = 0.2$, for the purpose of showing clearly the way in which the equilibrium distribution is upset by the motion of the dislocation. The results of this are given in figure 2 in the form of the variation of c/c_0 along the four lines $y_1 = +1, y_2 = +3.5, -y_1$ and $-y_2$ of figure 1 for the cases where the dislocation is at rest (full lines) and where it is moving (broken lines); x and y are here measured in

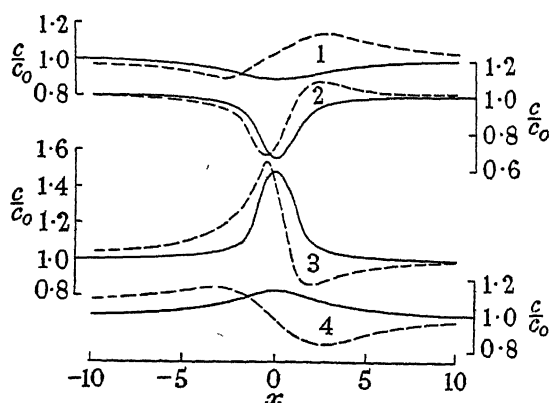


FIGURE 2. Variation of concentration of large solute atoms near a positive edge dislocation. Full lines, dislocation stationary; broken lines, dislocation moving slowly ($q = 0.2$) along positive x -axis. Curves: 1, $y_2 = +3.5$; 2, $y_1 = +1.0$; 3, $y_1 = -1.0$; 4, $y_2 = -3.5$.

terms of $\rho = 1$ as the unit of length. It will be seen that the main effects produced by the motion of the dislocation are the displacement of the centre of the distribution to the trailing side of the dislocation and the creation of a wave or trough on the leading side. These can be understood qualitatively by considering how an atmosphere

would adjust itself to a dislocation that was suddenly jerked forward through a small distance and then held fixed. The disturbance of the equilibrium distribution is smaller along those lines passing close to the centre of the dislocation than along the more distant ones; this is because the interaction forces, and hence the rate of relaxation, are greater near the centre of the dislocation.

5. THE PERTURBATION FORCE

The force exerted in the slip direction on a dislocation by a solute atom is $\partial V/\partial x$; hence, if μ is the number of solvent atoms per cm.³, the total force in the slip direction on unit length of the dislocation, from its atmosphere, is

$$F = \mu \iint c \frac{\partial V}{\partial x} dx dy.$$

If the concentration is written in the form $c = c_0(1 + Q)\exp(-V/kT)$, where Q is the relative perturbation at x, y , the Maxwellian term, $c_0\exp(-V/kT)$, can be omitted since it does not contribute to the force in the slip direction. Introducing the variables ρ and θ , and noting that $\partial V/\partial x = -A \sin 2\theta/r^2$, the force becomes

$$F = -A\mu c_0 \iint \exp(-V/kT) Q \frac{\sin 2\theta}{\rho} d\rho d\theta. \quad (15)$$

This expression has been integrated numerically by selecting a value for ρ and then evaluating the integrand at a series of annular points spaced at intervals of 15° , obtaining in this way the contribution of force from an annulus of the chosen radius. These contributions have been determined at $\rho = 0.2, 0.25, 0.5, 1, 2, 4, 5$ and 10 , using the values of the physical parameters discussed in the next section, and are assembled in figure 3. The area under this curve from $\rho = 0.1$ (corresponding to $r \simeq 5 \text{ \AA}$) outwards, when multiplied by $A\mu c_0$, gives the total perturbation force.

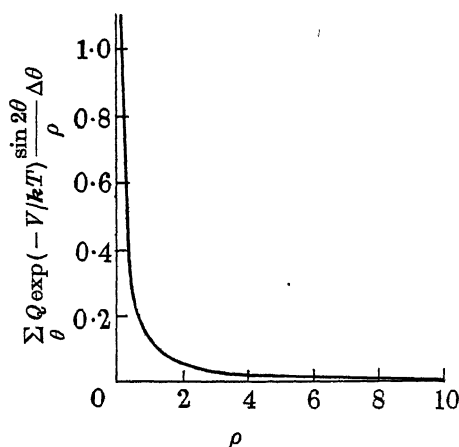


FIGURE 3

The approximately linear increase of this force with the speed of the dislocation, for small speeds, can be shown by examining the deviation of ψ , in equation (13),

from its equilibrium form corresponding to the case when the dislocation is at rest. When $\rho \leq 1$ and q is sufficiently small, the terms in equation (13) differ from their equilibrium values by $O(q)$ (Jaswon 1949) or by $O(\sqrt{q\rho})$. Hence ψ can be written in the form $[1 + O(q) + O(\sqrt{q\rho})] \exp(-V/2kT)$ or as $[1 + O(v)] \exp(-V/2kT)$ since $q = vl/4D$ and $\sqrt{q\rho} = vr/2D$. Noting also that $\sqrt{\phi} = [1 + O(\sqrt{q\rho})] \exp(-V/2kT)$ when $\rho \leq 1$, then

$$c/c_0 = [1 + O(v)] \exp(-V/kT). \quad (16)$$

The same result can also be deduced for the region where $\rho > 1$ by making use of the fact that the function $\exp(V/2kT) \exp(-vx/2D)$, like the function ψ , has the same value at the points ρ , θ and $1/\rho$, $-(\theta + \frac{1}{2}\pi)$. Since c/c_0 can be written in the above form for small v , then Q in equation (15), and hence the perturbation force, is linearly proportional to the speed of the dislocation.

6. CREEP DUE TO SLOW DISLOCATIONS

Because the speed of a slow dislocation is limited by the mobility of nearby solute atoms, examples of flow due to slow dislocations are best looked for in slow creep tests, of which those of Chalmers (1936) on tin crystals are particularly suitable. Micro-creep occurs in tin crystals at room temperature at rates of tensile strain below about $10^{-8} \text{ sec.}^{-1}$ and at tensile stresses below 150 g./mm.^2 . Above this range rapid flow occurs, typical of a plastic crystal stressed above its yield strength. The three main characteristics of micro-creep are:

- (i) the initial rate of creep increases linearly with the applied stress, being zero only at zero stress;
- (ii) under constant stress the rate of creep falls steadily with increasing strain, becoming practically zero after a total strain of 3×10^{-5} ;
- (iii) once micro-creep has been exhausted at any given stress, change to another stress within the micro-creep range gives no more creep.

These effects led Chalmers to suggest that micro-creep is caused by the slow migration of dislocations present at the start of the test, since in this case the creep would slow down and stop as these dislocations were used up. This is consistent with the recent view (Frank 1948) that only fast dislocations can multiply and so produce a large and rapid glide. If Chalmers is correct, the possibility arises that the dislocations move slowly because they are held back by atmospheres, in which case the first characteristic of micro-creep would be explained by the linear variation of the perturbation force with speed.

A quantitative comparison can be attempted by evaluating the perturbation force. Chalmers gave the tensile coefficient of the micro-creep rate as 3×10^{-9} per min., per g./mm.^2 . This corresponds to a shear rate coefficient of 2×10^{-15} per sec., per dyne/cm.^2 , since the multiplicity of glide planes in tin allows one to assume that slip occurs at 45° to the axis of tension. A shear stress τ on the glide planes in the glide direction produces a force $\tau\lambda$ per unit length on a dislocation (Mott & Nabarro 1948), and in a steady state of motion this force must be equal and opposite to F , the perturbation force. If the density of dislocations is $p \text{ cm.}^{-2}$ then the rate of shear

strain is $v\rho\lambda\text{sec.}^{-1}$. Hence, if the theory is correct the relation $2 \times 10^{-15} = v\rho\lambda^2/F$ should be obeyed. Substituting $q = v\lambda/4D$ gives this test equation as

$$2q\lambda^2\rho D/lF = 10^{-15}. \quad (17)$$

The calculation of l and F requires a knowledge of ϵ and c_0 defined in equations (1) and (2) respectively. The two grades of tin that Chalmers examined, one of which had been degassed, were observed to have the same initial creep rates. Although these samples were of different purities the total copper and iron content was about the same in each. Moreover, the degree of misfit, as determined from atomic diameters, is greater for these metals than for any of the other impurities that were reported present; taking the diameters for tin, iron and copper as 3.10, 2.48 and 2.55, respectively, $\epsilon = -0.20$ for iron and -0.18 for copper. It will therefore be assumed that these impurities formed the atmospheres and the corresponding values $\epsilon = -0.2$ and $c_0 = 3 \times 10^{-5}$ will be used. With $\lambda = 3 \times 10^{-8}$ cm., A and l are 9.2×10^{-21} dyne cm.² and 2.3×10^{-7} cm. respectively, and the area under the curve of figure 3 is 0.68 in non-dimensional units. With $\mu = 3.7 \times 10^{22}$ atoms/cm.³, the perturbation force is therefore $F = 0.68A\mu c_0 = 7 \times 10^{-3}$ dyne/cm. Substituting F and $q = 0.04$ in equation (17) then gives $\rho D = 2 \times 10^{-8}$ c.g.s. units. This result could be explained by the plausible assumptions of a density of dislocations of 10^7 to 10^8 cm.⁻², which is in the range usually suggested for soft single crystals, and a diffusion coefficient of 10^{-14} to 10^{-16} cm.²sec.⁻¹, which is reasonable for a low melting point metal at room temperature. No data on diffusion rates in tin appear to be available.

Two other numerical comparisons are possible. Only when q is small can the perturbation force increase linearly with the dislocation speed. For faster speeds, where q approaches unity, higher order terms become important and cause a deviation from linearity, leading eventually to the unstable range where the dislocations break away. The corresponding critical range of stress should therefore be given approximately by substituting $q = 1$ in equation (17); with $\rho D = 2 \times 10^{-8}$ this gives the critical tensile stress as 120 g./mm.² in agreement with observation. The critical creep rate can be examined by an independent method. The critical dislocation speed must be determined by the condition that solute atoms near a dislocation cannot keep up with it; it is therefore given in order of magnitude from equation (3) as $v \simeq (D/kT) |\nabla V| = Dl/r^2$, where r is some representative small distance from the centre of the dislocation. This corresponds to a critical rate of tensile creep of $\rho D \lambda l / 2r^2$; equating with Chalmers value of approximately 10^{-8} sec.⁻¹ gives $r \simeq 8 \text{ \AA}$, which is reasonable for the distance which might be taken when estimating the critical speed. Since $\rho^2 = v r^2 / Dl$ by definition, this method for the critical speed involves the condition $\rho = 1$; the alternative criterion $q = 1$ coincides with this when it is assumed that the representative value of r is $\frac{1}{2}l$.

It is thus evident that the type of creep expected from slow dislocations resembles micro-creep in several ways, within the limitations of existing data. A knowledge of the dependence of micro-creep on temperature and specific impurities, and of the diffusion rates of these impurities, is required before a really critical comparison can be made.

We wish to acknowledge the encouragement and support given to us by Professor D. Hanson, and the benefit we have had from several useful discussions with Professor N. F. Mott and Mr F. R. N. Nabarro, during the progress of this work. Financial help has been given to one of us (M.A.J.) by the award of a University Fellowship.

REFERENCES

- Chalmers, B. 1936 *Proc. Roy. Soc. A* **156**, 427.
 Copson, E. T. 1935 *Functions of a complex variable*. Oxford: Clarendon Press.
 Cottrell, A. H. 1948 *Report on the strength of solids*, p. 30. London: Physical Society.
 Cottrell, A. H. & Bilby, B. A. 1949 *Proc. Phys. Soc.* **62** (1), 49.
 Frank, F. C. 1948 *Report on the strength of solids*, p. 46. London: Physical Society.
 Jaswon, M. A. *Proc. Edinb. Math. Soc.* (in the Press).
 McLachlan, N. W. 1947 *Theory and application of Mathieu functions*. Oxford: Clarendon Press.
 Mott, N. F. & Nabarro, F. R. N. 1948 *Report of Bristol Conference on the strength of solids*, p. 1. London: Physical Society.
 Nabarro, F. R. N. 1947 *Proc. Phys. Soc.* **59**, 256.
 Nabarro, F. R. N. 1948 *Report on the strength of solids*, p. 38. London: Physical Society.
 Whittaker, E. T. & Watson, G. N. 1946 *Modern analysis*. Cambridge University Press.

The absorption and dispersion of ultrasonic waves in acetic acid

BY J. LAMB, *Electrical Engineering Department, Imperial College, London*
 AND J. M. M. PINKERTON, *Cavendish Laboratory, University of Cambridge*

(Communicated by Sir Lawrence Bragg, F.R.S.—Received 1 April 1949)

The paper gives the results of measurements of the absorption and velocity of propagation of ultrasonic waves in acetic acid over the frequency range 0.5 to 67.5 Mc./sec. and at temperatures from 16 to 60° C. It is shown that a dispersion and a maximum in the value of absorption per wave-length occur, and the results confirm the existence of a relaxation process arising from the perturbation of a molecular equilibrium by the ultrasonic vibrations. The absorption coefficient, α , at a frequency, ν , is represented by an equation of the form

$$\alpha = B\nu^2 + \frac{A\nu^2}{1 + (\nu/\nu_m)^2},$$

where A , B and ν_m are parameters independent of ν but varying with temperature, and ν_m is related to the relaxation time, τ , by $\nu_m = \frac{1}{2\pi\tau}$. These parameters are evaluated for a series of temperatures.

The results are discussed in the light of existing theory, and in particular the activation energy of the process is obtained from the measured variation of relaxation time with temperature. It is suggested that the relaxation mechanism is possibly connected with a perturbation of the equilibrium between single and double molecules of acetic acid. The results also indicate the existence of a further relaxation process giving a second maximum in the absorption per wave-length at some frequency greater than 67.5 Mc./sec.

1. INTRODUCTION

It is now well established (Pellam & Galt 1946; Rapuano 1947; Kittel 1946, 1948; Pinkerton 1949*a*) that the pressure absorption coefficient α relating to the propagation of ultrasonic waves through liquids, is in most cases appreciably greater than the value which can be attributed to viscous loss (Stokes 1845) and to thermal conductivity (Kirchhoff 1868). The accepted explanations of this discrepancy postulate the occurrence of a relaxation phenomenon arising from the perturbation by the ultrasonic wave of some molecular equilibrium. Any such relaxation mechanism would be expected to cause both dispersion and a maximum in the value of $\mu = \alpha\lambda$, the coefficient of absorption per wave-length, at a frequency related to the relaxation time. However, examination of the experimental results (Pinkerton 1949*a*) does not reveal much reliable evidence of these two effects.

It is the purpose of this paper to present results of new measurements of the absorption and velocity of propagation of ultrasonic waves in acetic acid over a wide range of frequencies (0.5 to 67.5 Mc./sec.) and temperatures (16 to 60° C), and to show that these results confirm both the dispersion and the maximum value of μ required by a relaxation process. Preliminary accounts of this work have already appeared (Pinkerton 1948; Lamb, Andreae & Bird 1948).

2. PREVIOUS WORK

The excess absorption in gases has been attributed (Kneser 1931, 1935) to the equilibrium between molecules having differing amounts of vibratory energy, and Kneser (1938) and Herzfeld (1941) considered that a similar explanation might also account for the increased absorption in some liquids. Hall (1948), following a suggestion by Herzfeld, ascribed the absorption in water to perturbation of the equilibrium between an ordered and a disordered arrangement of molecules in a lattice.

The measurements of Bazulin (1936) demonstrated that in acetic acid a maximum value of μ occurs at some frequency less than 3 Mc./sec. at 17° C, whilst a dispersion of about 1 % was claimed by Spakowski (1938) between 0.25 and 3 Mc./sec. The latter suggested that, in this case, the absorption might be due to the equilibrium between double and single molecules, on the grounds that a relaxation time of about 10^{-7} sec., indicated by his results, was much greater than that expected for the exchange of vibratory energy.

The theories of Stokes and Kirchhoff require α/ν^2 to be independent of frequency, ν , at a given temperature. For methyl and ethyl acetates and for solutions of certain metallic acetates several authors (Claeys, Errera & Sack 1937; Bär 1937; Bazulin 1939; Beyer & Smith 1946) have reported departures from this. These variations in α/ν^2 are, however, not accurately in accordance with the relaxation hypothesis, and, moreover, there is no record of a dispersion in these liquids.

3. EXPERIMENTAL METHOD

All the measurements were made by the pulse method which has already been fully discussed (Pellam & Galt 1946; Pinkerton 1949*b*); two complementary sets of observations were taken independently by the two authors of this paper, at the higher

frequencies (7.5 to 67.5 Mc./sec.) by J.M.M.P., and at the lower frequencies (0.5 to 9.8 Mc./sec.) by J.L. The frequency ranges of the two-series of measurements were thus arranged to overlap between 7 and 10 Mc./sec. The observations at the higher frequencies were taken in the Fresnel region of the X-cut quartz crystal generating the ultrasonic pulses, whilst at the lower frequencies the Fraunhofer region was used and a suitable correction applied to allow for the divergence of the ultrasonic beam, as previously explained (Born 1943; Pinkerton 1949*b*).

The absorption in acetic acid at the highest frequency was about 60 db./cm., so that a quartz bar had to be used as a delay line, in order that the reflected echo might be distinguished from the transmitted pulse when viewed on the screen of the cathode-ray tube (Rapuano 1947). At the lower frequencies the technique described by Pinkerton (1949*b*) was modified somewhat, in that at each frequency separate, but similar, quartz crystals were used to generate and detect the ultrasonic vibrations respectively. The high-frequency apparatus used for detecting the waves was thus separated electrically from the generating circuits. The latter consisted of a signal generator feeding a high-frequency amplifier, the push-pull output stage of which was biased well beyond cut-off. Thus the transmitting quartz crystal was only excited into vibration upon the receipt of a positive pulse of large amplitude on the grids of the two output valves. The received high-frequency pulses were amplified and displayed, after rectification, as uni-directional pulses on a cathode-ray tube, the time base of which controlled the pulse repetition frequency. The duration of each pulse was so chosen that it contained at least forty complete high-frequency oscillations. The receiving crystal was connected by a short length of low-capacity coaxial cable to a piston attenuator operated in the H_{11} mode (Gainsborough 1947), the attenuation constant of which was approximately 3 db./cm.; this was checked for linearity of operation at each frequency against a similar type of attenuator of larger diameter (attenuation constant = 2 db./cm.). At frequencies less than 2 Mc./sec. the insertion loss of the piston attenuator became excessive, and it was replaced by a calibrated resistance attenuator, fed from the receiving crystal through a suitable matching unit. In this way the input voltage of the sensitive receiving amplifier was constant during each measurement.

The velocity of ultrasonic waves in the liquid was measured directly by feeding a small continuous wave signal from the signal generator into the receiver and counting the number of beats of the recorded pulse as the receiving crystal was moved through a known distance.

In all the experiments the temperature of the liquid was carefully stabilized, particularly at the lower frequencies where the smaller absorption necessitated longer distances from source to detector.

Each frequency of operation was measured by comparison with the known harmonics of a standard quartz crystal oscillating at a frequency of 100 kc./sec.

4. EXPERIMENTAL RESULTS

Observations were taken over a range of temperatures (from 16 to 60° C) at fifteen fixed frequencies (0.5 to 67.5 Mc./sec.), which were selected to correspond roughly to equal increments in $\log \nu$. At each frequency the absorption coefficient α was found

directly from the observations, but the curves of α plotted against T corresponding to the lower frequencies lie so close together, that in order to present all these results on one diagram it is more convenient to plot $\log_{10}(\alpha/\nu^2)$ instead of α against T (figure 1a); at the higher frequencies the curves of α against T are well spaced (figure 1b).

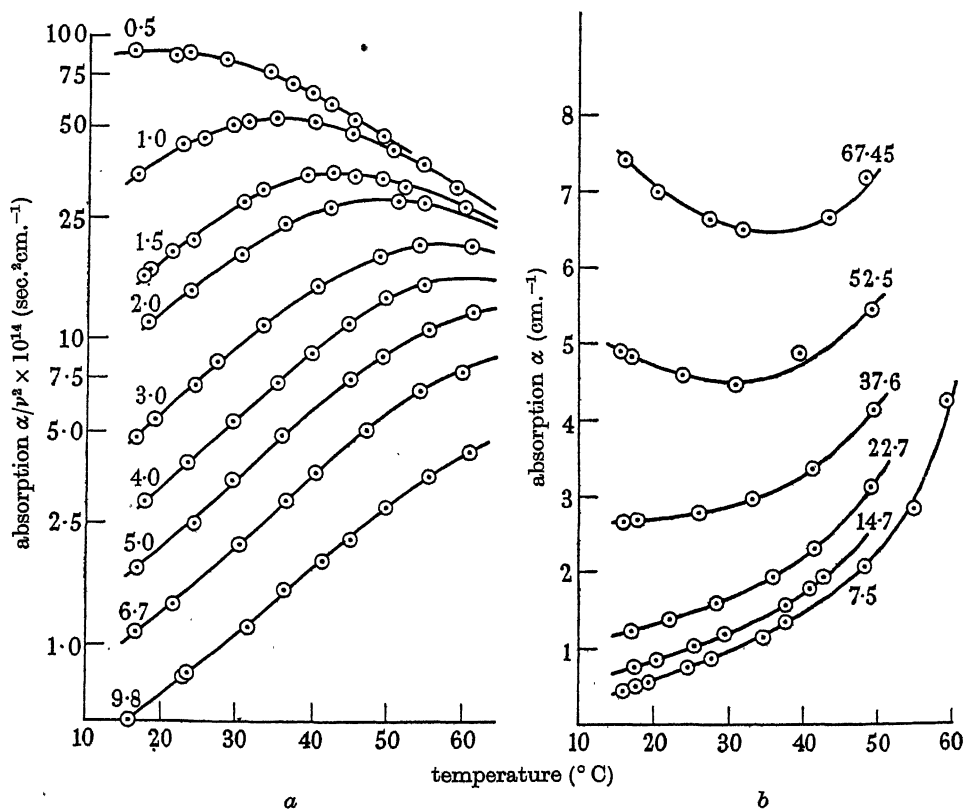


FIGURE 1. Absorption in acetic acid; measured values of absorption as a function of temperature. The figures against each curve denote the corresponding frequency of measurement in Mc./sec. The curves of α plotted against T corresponding to the lower frequencies lie so close together that it is here more convenient to plot $\log \alpha/\nu^2$ against T . At the higher frequencies the curves of α against T are well spaced.

The velocity of propagation was measured at each of the frequencies used between 0.5 and 9.8 Mc./sec., but the results for 0.5, 4 and 5 Mc./sec. only are given in figure 2. The values of the velocity at frequencies above 4 Mc./sec. coincide with the upper line, whilst the lower line gives the corresponding values for the lowest frequency, 0.5 Mc./sec.; the experimental points for the intervening frequencies lie between the two lines.

The measured values of absorption and velocity did not appear to be affected by the purity of the acetic acid; samples graded Analar and the ordinary commercial glacial acetic acid gave similar results. All samples were rejected before corrosion of the container had contaminated them sufficiently to affect the observations.

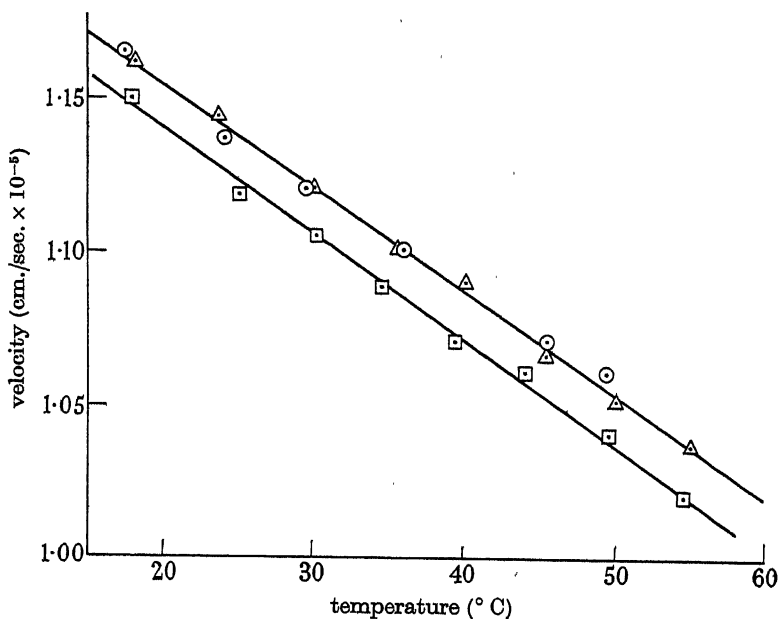


FIGURE 2. Dispersion in acetic acid. The velocity values at frequencies above 4 Mc./sec. all coincide with the upper line but only points for 4 and 5 Mc./sec. are shown. The lower line refers to the lowest frequency of measurement, 0.5 Mc./sec. Experimental points for frequencies between 0.5 and 4 Mc./sec. lie between the two lines but are not shown. ○ 5 Mc./sec., △ 4 Mc./sec., □ 0.5 Mc./sec.

5. THEORETICAL ANALYSIS OF THE EXPERIMENTAL RESULTS

It will be shown in § 7 below that if the absorption can be ascribed to a single relaxation mechanism in addition to the normal contributions due to viscosity and thermal conductivity, then it may be represented to a good approximation by

$$\alpha = B\nu^2 + \frac{A\nu^2}{1 + (\nu/\nu_m)^2} \quad (1)$$

Here A , B and ν_m are parameters independent of ν but varying with temperature and $\nu_m = \frac{1}{2\pi\tau}$, where τ is the relaxation time. It is noteworthy that the absorption per wave-length corresponding to the second term on the right-hand side of (1) is similar in form to that given by Debye (1929) for the variation of loss factor with frequency in the case of the dielectric absorption associated with dipole relaxation.

The above parameters were obtained for a series of temperatures between 16 and 60°C by the methods given in § 6 and are listed in table 1. Using these values curves of α/ν^2 as a function of ν have been evaluated and are shown in figure 3 for three temperatures only. Comparison of these curves with the experimental points, also given in figure 3, confirms that equation (1) applies for acetic acid over the frequency range of the observations.

A more familiar representation may be given to the experimental results by considering the absorption per wave-length, μ' , due to the relaxation mechanism alone,

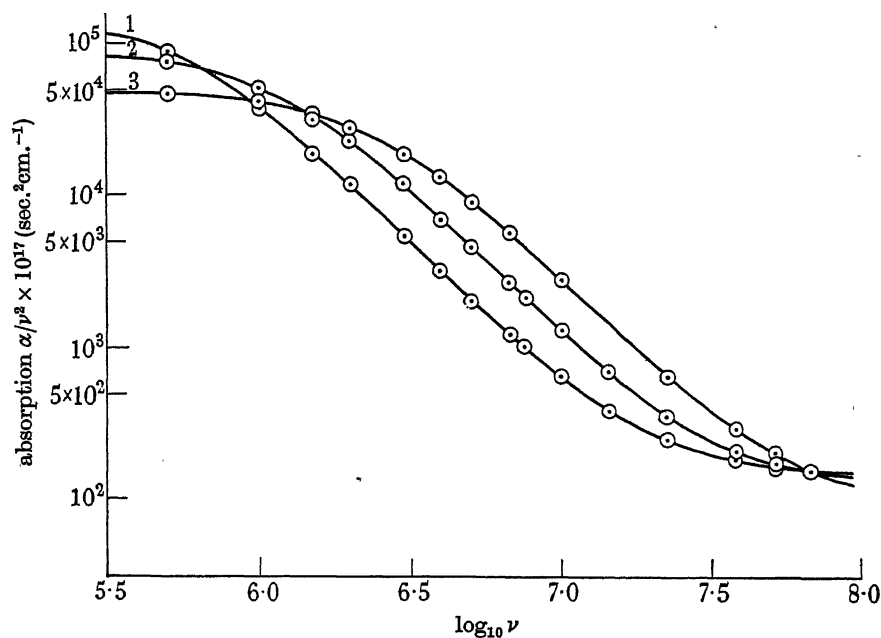


FIGURE 3. α/ν^2 as a function of frequency. The curves are calculated using the values given in table 1, and the points are taken from the smooth curves of figure 1. Curve 1, 20° C.; 2, 35° C.; 3, 50° C.

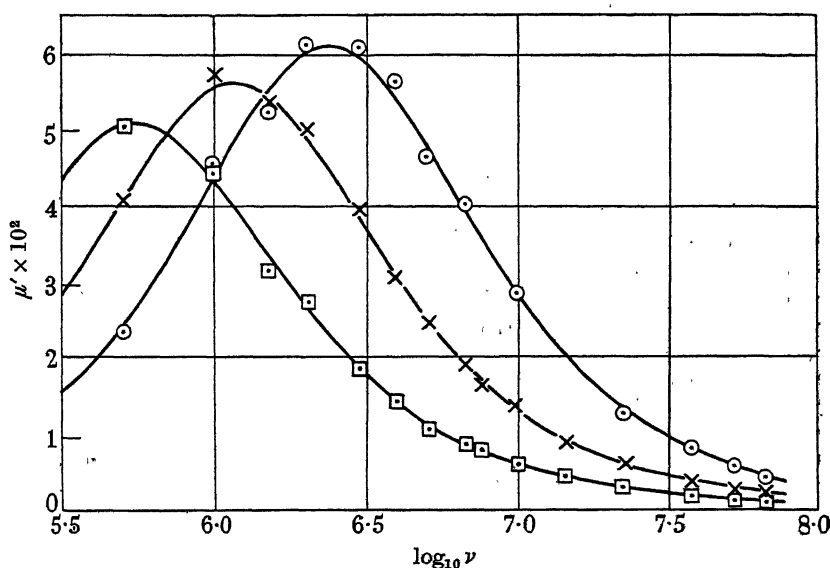


FIGURE 4. $\mu' = \alpha'V/\nu$, the absorption per wave-length due to the relaxation process alone, as a function of frequency. The curves are calculated using the values given in table 1. Values of α have been taken from the smooth curves of figure 1 to give $\alpha' = \alpha - B\nu^2$. $\odot$ 50° C, $\times$ 35° C, $\square$ 20° C.

as a function of frequency. Putting $\alpha' = \alpha - B\nu^2$ then $\mu' = \alpha'\lambda = \alpha'V/\nu$, which may also be calculated from the parameters of table 1. Figure 4, in which the curves have been calculated and the points represent the observations, shows the variation of μ' with ν .

6. EVALUATION OF THE PARAMETERS A , B AND ν_m

It has been suggested that in cases where some of the absorption is due to a relaxation and some due to viscosity and thermal conductivity a relation of the form of (1) should give the dependence of α upon ν . A test of this equation, using the experimental results, requires the choice of the best values for A , B and ν_m to represent the absorption coefficient observed at fifteen different frequencies between 0.5 and 67.5 Mc./sec. This fit must be made with different values of the parameters at each temperature considered and has been done in the following way. When $\nu \gg \nu_m$ equation (1) becomes approximately

$$\alpha \approx B\nu^2 + A\nu_m^2, \quad (2)$$

which, since $A\nu_m^2$ is independent of ν , may be tested by plotting α against ν^2 . This plot for three temperatures is given in figure 5, in which the values of α have been taken from the smooth curves of figure 1*b*. Figure 5 shows that equation (2) is accurately obeyed at all the temperatures for three frequencies between 22.7 and 52.5 Mc./sec.; the departure of the points from the straight lines at lower frequencies is due to the fact that the condition $\nu \gg \nu_m$ is no longer satisfied. A similar discrepancy in the points corresponding to 67.5 Mc./sec. is discussed in § 8. The value of the constant B in equation (2) is deduced at each temperature from graphs similar to those of figure 5 and is given in table 1.

TABLE 1

tem- perature (°C)	$B \times 10^{17}$ (sec. ² cm. ⁻¹)	$A \times 10^{17}$ (sec. ² cm. ⁻¹)	ν_m (Mc./sec.)		$V_{\min.} \times 10^{-5}$ or $V_0 \times 10^{-5}$ (cm. sec. ⁻¹) $\nu = 0.5$ Mc./sec.	$V_{\max.} \times 10^{-5}$ or $V_{\infty} \times 10^{-5}$ (cm. sec. ⁻¹) $\nu \geq 4$ Mc./sec.	$\mu'_{\max.} \times 10^2$ $= \frac{AV\nu_m}{2}$
			ν_{m1} first estimate	ν_{m2} second estimate			
17	160(0)	178,000	0.493 ₀	0.488 ₀	1.15 ₂	1.16 ₈	5.01
20	151(5)	160,000	0.552 ₇	0.556 ₀	1.14 ₁	1.15 ₈	5.10
25	140(1)	132,000	0.704 ₉	0.725 ₀	1.12 ₃	1.13 ₈	5.37
30	132(0)	107,000	0.910 ₁	0.915 ₀	1.10 ₇	1.12 ₂	5.45
35	124(0)	88,700	1.16 ₇	1.15 ₈	1.08 ₉	1.10 ₄	5.61
40	118(6)	72,900	1.46 ₈	1.48 ₉	1.07 ₁	1.08 ₇	5.83
45	112(5)	58,800	1.88 ₈	1.87 ₄	1.05 ₃	1.07 ₀	5.87
50	106(8)	48,400	2.37 ₂	2.40 ₈	1.03 ₈	1.05 ₃	6.10
55	100(8)	41,100	3.00 ₇	2.98 ₅	1.01 ₉	1.03 ₈	6.35
60	95(0)	33,800	3.78 ₅	3.77 ₂	1.00 ₁	1.01 ₉	6.41

* The value of V used here is that corresponding to a frequency ν_{m2} at the temperature considered.

Once B is known values for A and ν_m are found by writing $\alpha' = \alpha - B\nu^2$ in equation (1) which may then be transformed into

$$\frac{1}{\alpha'} = \frac{1}{A} \left[\frac{1}{\nu^2} + \frac{1}{\nu_m^2} \right]. \quad (3)$$

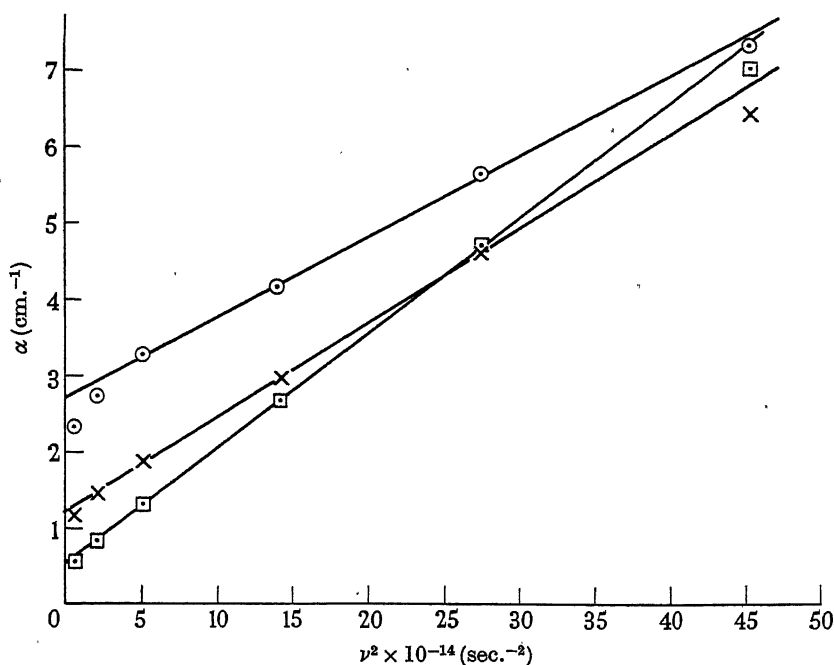


FIGURE 5. α plotted against ν^2 for $\nu > \nu_m$. As explained in the text, the straight lines have been drawn through the points at three frequencies only, namely at 22.7, 37.6 and 52.5 Mc./sec. for which $\nu \gg \nu_m$. The parameter B , at any temperature, is given by the slope of the corresponding straight line. ○ 50° C, × 35° C, ◻ 20° C.

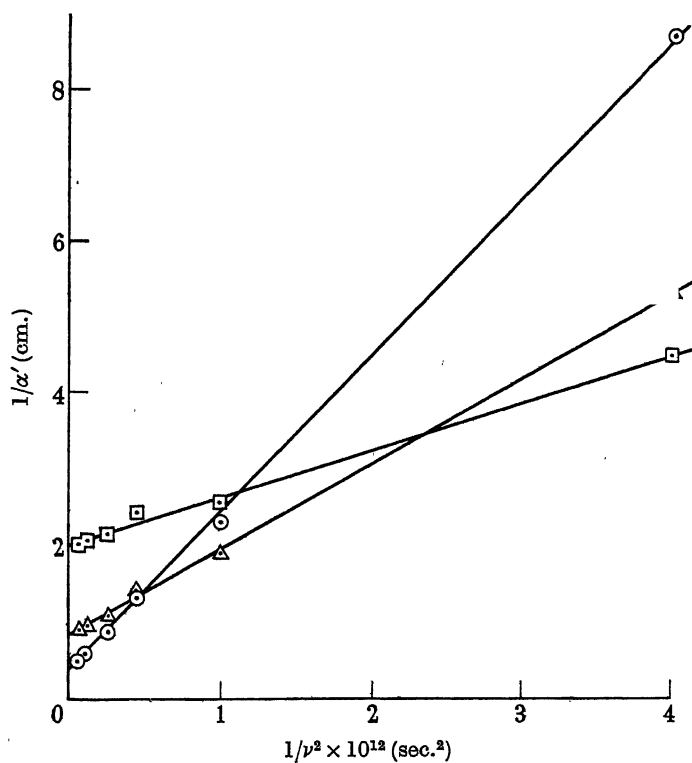


FIGURE 6. Linear relationship between $1/\alpha'$ and $1/\nu^2$ at different temperatures. The slope gives the parameter A whilst ν_{m1} is obtained from the intercept on the $1/\alpha'$ axis. ○ 50° C, △ 35° C, ◻ 20° C.

A plot of $1/\alpha'$ against $1/\nu^2$ should therefore be a straight line, which can be used to find A and ν_m . Typical plots are shown in figure 6, and the corresponding values of A and ν_m are given in columns 3 and 4 of table 1. These estimates of ν_m , denoted ν_{m1} , may not be the most accurate, since in drawing the lines of figure 6 equal weight cannot readily be given to each point. If the value of A is regarded as known, a second estimate, ν_{m2} , is possible by again transforming (3), to give

$$\sqrt{\left(\frac{A}{\alpha'/\nu^2} - 1\right)} = \nu/\nu_m. \quad (4)$$

A plot of $\sqrt{\left(\frac{A}{\alpha'/\nu^2} - 1\right)}$ against ν should be a straight line passing through the origin, and the points are more uniformly spaced. These plots are shown in figure 7 and the corresponding values of ν_{m2} in column 5 of table 1.

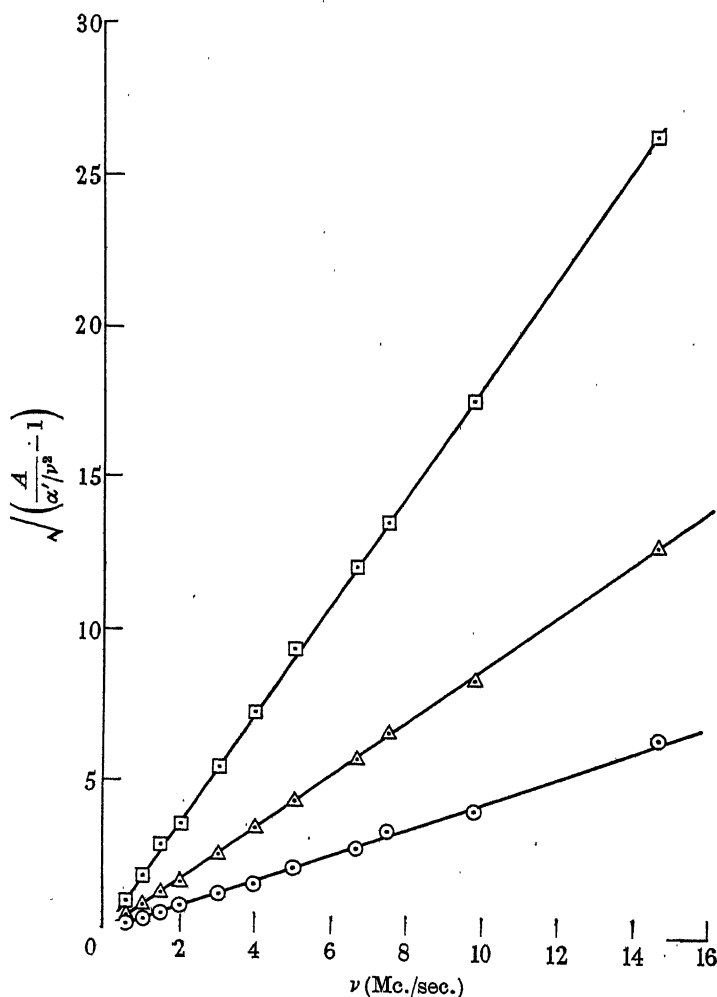


FIGURE 7. Linear relationship between $\sqrt{\left(\frac{A}{\alpha'/\nu^2} - 1\right)}$ and ν at different temperatures. The slope gives the parameter ν_{m2} , the more reliable estimate of $\nu_m = \frac{1}{2\pi\tau}$. ○ 50° C, △ 35° C, □ 20° C.

7. DETAILED THEORY

The following treatment is based on that used by Kneser (1931, 1938), Rütgers (1933), Richards (1939), Herzfeld (1941) and Bauer (1949). In view of what has previously been written on this subject only an outline is necessary here, in order to reveal certain obscurities and to consider the variation of the relaxation time with temperature.

In gases the absorption is discussed in terms of an equilibrium of the form



where A_2 and A_1 represent states of the same molecule differing by one quantum of vibrational energy. An equilibrium of the same form is assumed to apply for liquids, but it is not necessary to specify initially the nature of the energy change, ΔE . If n_1, n_2 are the molar fractions in states A_1, A_2 respectively, and K_1, K_2 the corresponding rate constants, dependent solely on the temperature, then the kinetic relation is

$$\frac{dn_2}{dt} = K_1 n_1 - K_2 n_2. \quad (5)$$

By definition $n_1 + n_2 = 1$, and at equilibrium $K_1 n_1 = K_2 n_2$ and also $n_2/n_1 = e^{-\Delta E/RT}$ according to the Boltzmann distribution law.

A small perturbation of the equilibrium condition varying sinusoidally with time at frequency $\omega/2\pi$ may be shown (e.g. Hall 1948) to produce a change δn_2 in n_2 given by

$$\delta n_2 = \frac{-\delta[\Delta E/RT]}{2[1 + \cosh(\Delta E/RT)]} \frac{1}{1 + j\omega/(K_1 + K_2)}, \quad (6)$$

where $j = \sqrt{-1}$ and the factor $e^{j\omega t}$ has been omitted. Equation (6) involves no approximation except that $(\delta n_2/n_2)^2$ and higher powers are neglected, as is entirely justified for the small disturbance caused by the ultrasonic wave. The dependence of n_2 on frequency may be characterized by a relaxation time $\tau = 1/(K_1 + K_2)$.

It is now necessary to consider the mechanism by which the equilibrium is disturbed. For water Hall (1948) supposed that it is perturbed by the ultrasonic changes in pressure, and consequently evaluated $[\partial n_2/\partial p]_T$. In the present case it is assumed that, as in gases, there is no change in volume associated with the transition from A_1 to A_2 ; the equilibrium is assumed to be perturbed by a change in temperature only. If ΔE is independent of T , equation (6) gives

$$\frac{\delta n_2}{\delta T} = \left[\frac{\partial n_2}{\partial T} \right]_p = \frac{(\Delta E/RT^2)}{2[1 + \cosh(\Delta E/RT)]} \frac{1}{(1 + j\omega\tau)}. \quad (7)$$

The quantity $C^1 = \Delta E[\partial n_2/\partial T]_p$ has the dimension of a specific heat, and represents the additional contribution to the specific heat of the system due to the perturbation. Since C^1 depends on ω it is usually called the relaxing specific heat; it is convenient to write $C^1 = C_i/(1 + j\omega\tau)$. Then from equation (7):

$$C_i = R \left[\frac{\Delta E}{RT} \right]^2 / 2 \left[1 + \cosh \left(\frac{\Delta E}{RT} \right) \right]. \quad (8)$$

If $\Delta E/RT > 4.5$, then to within 2 %

$$C_i = R \left[\frac{\Delta E}{RT} \right]^2 e^{-\Delta E/RT}, \quad (9)$$

which is the form usually quoted (e.g. Richards 1939).

Equation (8) shows that $C_i = 0$ when $\Delta E/RT = 0$, or $\Delta E/RT = \infty$, and takes a maximum value ($= 0.44R$) when $\Delta E/RT = 2.4$. Thus, equilibria involving either very small or very large energy changes give rise to small relaxing specific heats (and, as will be seen, to relatively small absorptions of ultrasonic waves). Now C_p is about $10R$ to $15R$ for most liquids, and if the approximate equation (9) applies C_i/C_p will be a small fraction and, in what follows, it may be neglected in comparison with unity.

More generally ΔE represents the change in free energy and must be written $\Delta F = \Delta E_0 - T\Delta S_0$, where ΔE_0 and ΔS_0 are the changes in internal energy and entropy respectively. Equation (7) then has to be written

$$\left[\frac{\partial n_2}{\partial T} \right]_p = \frac{[\Delta E_0/RT^2]}{2[1 + \cosh(\Delta F/RT)]} \frac{1}{(1 + j\omega\tau)},$$

and leads, when $\Delta F/RT > 4.5$, to a modification of equation (9), namely,

$$C_i = R \left[\frac{\Delta E_0}{RT} \right]^2 e^{-\Delta E_0/RT} e^{\Delta S_0/R}, \quad (9a)$$

in which the relaxing specific heat has been increased by the factor $e^{\Delta S_0/R}$.

The absorption and dispersion caused by the relaxing specific heat may be found, following Herzfeld (1941), by writing the complex velocity of sound

$$V^2 = \left[\frac{\partial p}{\partial \rho} \right]_{\text{adiabatic}}. \quad (10)$$

The adiabatic change is equivalent to one at constant entropy, and hence by a well-known thermodynamic relation

$$\left[\frac{\partial p}{\partial \rho} \right]_{\text{adiabatic}} = \left[\frac{\partial p}{\partial \rho} \right]_S = \gamma \left[\frac{\partial p}{\partial \rho} \right]_T, \quad (11)$$

where $\gamma = C_p/C_v$. These assumptions are not justified *a priori*, since the perturbation of the equilibrium implies an irreversible change; the only justification that can be offered here is that the relations deduced from them appear to give a satisfactory representation of the experimental results. It is assumed in what follows that $[\partial p/\partial \rho]_T$ does not depend on ω , which is equivalent to the previous assumption that $[\partial n_2/\partial p]_T$ may be neglected. C_p and C_v in (11) are regarded as including C^1 which depends on ω . $[\partial p/\partial \rho]_T$ is then eliminated from (10) and (11) by writing

$$V_0^2 = \frac{C_{p0}}{C_{v0}} \left[\frac{\partial p}{\partial \rho} \right]_T,$$

where V_0 is the velocity at very low frequencies and C_{p0} and C_{v0} are the specific heats normally measured. At any higher frequency $\nu = \omega/2\pi$

$$\frac{V^2}{V_0^2} = \frac{C_{p\omega}}{C_{v\omega}} \frac{C_{v0}}{C_{p0}}. \quad (12)$$

Herzfeld writes $C_{p\omega} = (C_{p\omega} - C_{v\omega}) + C_{v\omega} = \Delta + C_{v\omega}$, where Δ is supposed independent of ω , and thus $\Delta = C_{p0} - C_{v0}$. He also puts $C_{v\omega} = [C_{v\infty} + C_i/(1 + j\omega\tau)]$, where $C_{v\infty}$ is C_v at a frequency such that $\omega\tau \gg 1$.

Now the expression $p = p_0 e^{-\alpha'x} e^{j\omega[t - x/V]}$ represents an ultrasonic wave propagated in the x direction with velocity V_0 and coefficient of absorption α' . Hence the complex velocity V is given by $1/V = (1/V_0 - j\alpha'/\omega)$, and $V^2/V_0^2 \approx (1 + 2j\alpha'V_0/\omega)$ if $(\alpha'V_0/\omega)^2$ and higher powers are neglected.

Substitution for $C_{p\omega}$ and $C_{v\omega}$ in (12) may thus be shown to lead to a value of α'/ν^2 given by

$$\frac{\alpha'}{\nu^2} = \frac{2\pi^2}{V_0} (\gamma - 1) \frac{C_i}{C_{p0}} \frac{\tau}{(1 + \omega^2\tau)}, \quad (13)$$

assuming that $C_{v\infty}/C_{v0} \approx 1$. Following the notation of Kneser (1938), put $\tau = \frac{1}{2\pi\nu_m}$ so that (13) becomes

$$\frac{\alpha'}{\nu^2} = \frac{\pi}{V_0} (\gamma - 1) \frac{C_i}{C_{p0}} \frac{1}{\nu_m} \frac{1}{[1 + (\nu/\nu_m)^2]}. \quad (13a)$$

The absorption per wave-length $\mu' = \alpha'\lambda = \alpha'V/\nu$, and neglecting the small variation of V with ν , this is given by

$$\mu' = \pi(\gamma - 1) \frac{C_i}{C_{p0}} \frac{(\nu/\nu_m)}{[1 + (\nu/\nu_m)^2]}, \quad (14)$$

which shows that when $\nu = \nu_m$, μ' has the maximum value

$$\mu'_{\max.} = \frac{1}{2}\pi(\gamma - 1) \frac{C_i}{C_{p0}}. \quad (15)$$

If the appropriate expression for C_i be inserted in this relation from (8), (9) or (9a), there follow complete expressions for the absorption in terms of the parameters ΔE_0 , ΔS_0 , K_1 and K_2 determining the equilibrium, and the known quantities γ and C_{p0} .

Equation (12) may also be used to show that the relaxing specific heat causes the velocity to change with frequency. This dispersion has been assumed to be small in order to write a simple expression for α'/ν^2 ; Kneser (1938) has shown that the total change in velocity associated with a relaxation process is related to the absorption by

$$\frac{\Delta V}{V} = \frac{\mu'_{\max.}}{\pi}, \quad (16)$$

provided $\Delta V/V \ll 1$.

It will be of interest to consider the dependence of τ and therefore ν_m on the absolute temperature T . Now $\nu_m = \frac{1}{2\pi\tau} = (K_1 + K_2)/2\pi$, and if $\Delta E/RT > 4.5$ then $K_1/K_2 = e^{-\Delta E/RT} < 0.01$, so that K_1 may be neglected in comparison with K_2 . If the equilibrium between A_1 and A_2 is chemical in the broadest sense then K_2 might be expected to vary with T as required by

$$K_2 = DT^m e^{-\Delta E_a/RT}, \quad (17)$$

where D is a constant, m is zero or a small positive number, and ΔE_a is the apparent activation energy of the back reaction ($2 \rightarrow 1$); the activation energy of the forward reaction is $(\Delta E + \Delta E_a)$. Thus

$$\nu_m = \frac{K_2}{2\pi} = \frac{D}{2\pi} T^m e^{-\Delta E_a/RT}. \quad (18)$$

It will be possible to test this relation using the experimental results (see § 8).

If the possible occurrence of compressional viscosity be ignored the contribution of ordinary viscous loss to the absorption is given by Stokes' formula

$$\frac{\alpha_{\text{viscous}}}{\nu^2} = \frac{8\pi^2\eta}{3\rho V^3},$$

where η is the shear viscosity and ρ the density. Experimental values of α/ν^2 often exceed, but are never less than, the value given by this formula. It is reasonable to assume therefore that the viscous absorption always forms part of that measured, and must be suitably combined with the absorption due to relaxation. If each separate contribution to α satisfies the condition $\mu = \alpha\lambda \ll 1$, then it may readily be shown that the values are additive. Further, it is possible that more than one relaxation process may contribute to the loss at any given frequency. If, however, it be supposed that ν_m is the lowest relaxation frequency, and that ν'_m, ν''_m , etc., all greatly exceed ν_m , then, at frequencies near to or less than ν_m , these additional relaxation processes may be seen from (13a) to contribute terms to α varying as ν^2 . The effect of these other relaxations may therefore be included together with viscous absorption,* in a single term $B\nu^2$, and using (13a) the overall dependence of α on ν will be of the form previously quoted in § 5, namely

$$\alpha = B\nu^2 + \frac{A\nu^2}{1 + (\nu/\nu_m)^2}, \quad (1)$$

where A , B and ν_m are independent of ν but vary with T .

8. DISCUSSION

The agreement of equation (1) with experiment indicated by figures 3 to 7, which has been discussed in §§ 5 and 6, is regarded as conclusive evidence for the occurrence of a relaxation phenomenon in acetic acid. Unfortunately, the nature of the molecular equilibrium which is perturbed by the ultrasonic wave and gives rise to this relaxation is unknown from *a priori* considerations; it is not possible therefore to estimate the absorption directly from a previous knowledge of the parameters of the equilibrium $A_1 \rightleftharpoons A_2$.

The dependence of absorption α on ν has been shown to satisfy the relaxation hypothesis; it is of interest to see whether the dispersion is of the right order of magnitude, by relating it to the maximum value of μ' using equation (16), $\Delta V/V = \mu'_{\text{max}}/\pi$. At 30° C, for example, the measured change in velocity $\Delta V/V \approx 1.4\%$, whilst from the observed value of μ'_{max} , the calculated dispersion is 1.7%. This agreement is

* The loss due to thermal conductivity (Kirchhoff 1868) is negligible for most liquids.

within the experimental error for the determination of V , estimated to be $\frac{1}{4}\%$. At the lower temperatures the range of frequencies used was not quite sufficient to cover the total dispersion.

A further test of theory is possible from an examination of the dependence of ν_m and $\mu'_{\max}$ on the temperature T . The values of ν_{m2} in table 1 may be used to test the validity of equation (18). Thus if $m = 0$, a plot of $\log \nu_m$ against $1/T$ should yield

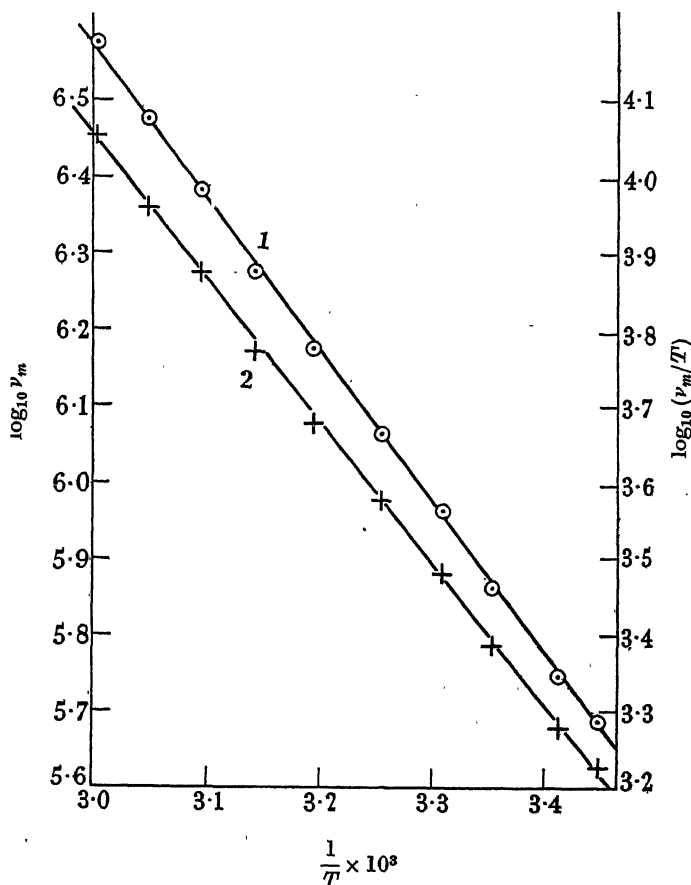


FIGURE 8. Dependence of $\log_{10} \nu_m$ and $\log_{10} (\nu_m/T)$ on $1/T$. The slope gives the apparent activation energy ΔE_a of the backward reaction ($2 \rightarrow 1$). From the plot of $\log_{10} (\nu_m)$ against $1/T$, curve 1, $\Delta E_a = 8.86$ kcal./mole, whilst from the plot of $\log_{10} (\nu_m/T)$ against $1/T$, curve 2, $\Delta E_a = 8.46$ kcal./mole.

a straight line, whereas if $m = 1$, $\log (\nu_m/T)$ should be a linear function of $1/T$. In fact, as shown by figure 8, the experimental values lie close to a straight line for both plots, and accordingly a specific value cannot be assigned to m . From the first plot of $\log \nu_m$ against $1/T$ the apparent activation energy, ΔE_a , of the backward reaction is 8.86 kcal./mole; the corresponding value from the plot of $\log (\nu_m/T)$ against $1/T$ is 8.46 kcal./mole.*

* The linear plots of figure 8 indicate that the activation energy is sensibly independent of temperature.

The dependence of $\mu'_{\max}$ on T may be compared with theory in an analogous way by substituting an expression for C_i in equation (15). Now $\Delta F = \Delta E_0 - T \cdot \Delta S_0$ and provided $\Delta F/RT > 4.5$, C_i is given by equation (9a), so that equation (15) gives

$$\mu'_{\max} = \frac{\pi(\gamma-1)R}{2C_{p0}} \left[\frac{\Delta E_0}{RT} \right]^2 e^{-\Delta E_0/RT} e^{\Delta S_0/R}. \quad (19)$$

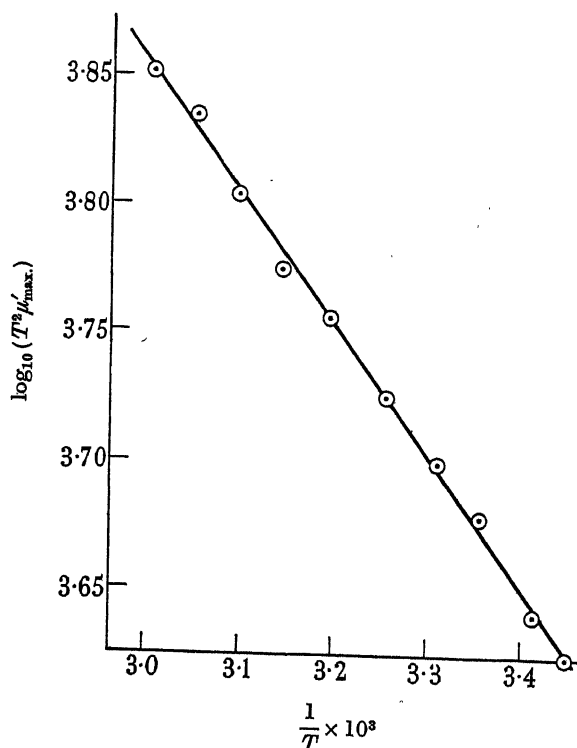


FIGURE 9. $\log(T^2 \mu'_{\max})$ plotted against $1/T$. The slope gives the difference in energy level between the states 1 and 2. The points correspond to the values of $\mu'_{\max}$ given in column 8 of table 1.

It has already been assumed that ΔS_0 is independent of T (an assumption which is not necessarily correct) and provided also that γ and C_{p0} can be regarded as independent of T , then $\log(T^2 \mu'_{\max})$ should be a linear function of $1/T$. When the values of $\mu'_{\max}$ (table 1) are plotted in this way the graph is a straight line as shown by figure 9. The slope of this line corresponds to a value of ΔE_0 of 2.32 kcal./mole. Now taking $\gamma = 1.22$, $C_{p0} = 28$ cal./mole/degree and $\mu'_{\max} = 5.45 \times 10^{-2}$, at 30°C , the value of C_i given by (15) is 4.41 cal./mole/degree and substitution of this in (9a) yields $\Delta S_0 = 3.89$ cal./mole/degree. Thus $\Delta F = 1.14$ kcal./mole and $\Delta F/RT = 1.89$; this, however, invalidates the condition for equation (19), that $\Delta F/RT > 4.5$, and consequently the energies derived from this equation are very approximate. Unfortunately, the more rigorous expression for C_i does not lend itself to a graphical examination of the experimental results.

Acetic acid is known to form double molecules united by two hydrogen bonds in the vapour phase, and this association is confirmed for the liquid by partition experiments between inert solvents, measurements of the permittivity of dilute solutions and electron diffraction experiments (Latimer & Rodebush 1920; Pauling & Brockway 1934; Broughton 1934; le Fèvre & Vine 1938; Halford 1941). It has been suggested that the absorption of ultrasonic waves in this liquid (Spakowski 1938; Richards 1939) arises from the perturbation of the equilibrium between single and double molecules, and it may well be that this explanation accounts for the present results. The heat of complete dissociation of the double molecules in the gas phase is 15.94 kcal./mole at 25° C (Halford 1941), but this value would not necessarily be the same in the liquid phase. The discrepancy between this value, however, and that obtained above suggests that the dissociation may proceed in two or more stages and that the relaxation mechanism observed is associated with the breakage of only one hydrogen bond. In this connexion it is noteworthy that the parameter B of equation (1) is much larger than is to be expected from viscosity alone; for example at 30° C, $\alpha_{\text{viscous}}/\nu^2 = 20 \times 10^{-17} \text{ sec.}^2 \text{ cm.}^{-1}$, whereas $B = 132 \times 10^{-17} \text{ sec.}^2 \text{ cm.}^{-1}$. It seems likely that the term $B\nu^2$ includes contributions to the absorption arising from at least one further relaxation process with ν_m equal to some frequency greater than 67.5 Mc./sec. This is also suggested by the fact that at 67.5 Mc./sec. the variation of α with temperature exhibits a minimum as shown in figure 1*b*. Moreover, in figure 5, the values of α corresponding to a frequency of 67.5 Mc./sec. lie below the straight lines drawn, particularly at the lower temperatures.

Possibly this second relaxation time might also be linked with the dissociation of the double molecule.

The authors wish to thank Mr J. A. Ratcliffe and Professor Willis Jackson for the close interest which they have shown in the present work; they are also grateful to Mr J. H. Andreae and Mr R. Bird for the assistance which they have given with some of the experimental work and to Dr A. R. Miller and Mr E. Bauer for many suggestions in discussion.

REFERENCES

- Bär, R. 1937 *Helv. phys. Acta*, **10**, 332.
 Bauer, E. 1949 *Proc. Phys. Soc. A*, **62**, 141.
 Bazulin, P. 1936 *C.R. Acad. Sci. U.R.S.S.* **3**, 285.
 Bazulin, P. 1939 *J. Phys. U.R.S.S.* **1**, 431.
 Beyer, R. T. & Smith, M. C. 1946 *J. Acoust. Soc. Amer.* **18**, 424.
 Born, H. 1943 *Z. Phys.* **120**, 383.
 Broughton, G. 1934 *Trans. Faraday Soc.* **30**, 367.
 Claeys, J., Errera, J. & Sack, H. 1937 *Trans. Faraday Soc.* **33**, 136.
 Debye, P. 1929 *Polar molecules*. New York: Chemical Catalog Co.
 Gainsborough, G. F. 1947 *J. Instn Elect. Engrs*, **94**, pt. III, 203.
 Halford, J. O. 1941 *J. Chem. Phys.* **9**, 859.
 Hall, L. 1948 *Phys. Rev.* **73**, 775.
 Herzfeld, K. F. 1941 *J. Acoust. Soc. Amer.* **13**, 32.
 Kirchhoff, G. G. 1868 *Ann. Phys., Lpz.*, **134**, 177.
 Kittel, C. 1946 *J. Chem. Phys.* **14**, 614.
 Kittel, C. 1948 *Rep. Progr. Phys.* p. 205.
 Kneser, H. O. 1931 *Ann. Phys., Lpz.*, **11**, 761.

- Kneser, H. O. 1935 *Z. tech. Phys.* **16**, 213.
 Kneser, H. O. 1938 *Ann. Phys., Lpz.*, **32**, 277.
 Lamb, J., Andreae, J. H. & Bird, R. 1948 *Nature*, **162**, 993.
 Latimer, W. M. & Rodebush, W. H. 1920 *J. Amer. Chem. Soc.* **42**, 1419.
 le Fèvre, R. J. W. & Vine, H. 1938 *J. Chem. Soc. (II)*, p. 1795.
 Pauling, L. & Brockway, L. O. 1934 *Proc. Nat. Acad. Sci., Wash.*, **20**, 336.
 Pellam, J. R. & Galt, J. K. 1946 *J. Chem. Phys.* **14**, 608.
 Pinkerton, J. M. M. 1948 *Nature*, **162**, 106.
 Pinkerton, J. M. M. 1949*a* *Proc. Phys. Soc. B*, **62**, 129.
 Pinkerton, J. M. M. 1949*b* *Proc. Phys. Soc. B*, **62**, 286.
 Rapuano, R. A. 1947 *Phys. Rev.* **72**, 78.
 Richards, W. T. 1939 *Rev. Mod. Phys.* **11**, 36.
 Rütgers, A. J. 1933 *Ann. Phys., Lpz.*, **16**, 350.
 Spakowski, B. 1938 *C.R. Acad. Sci. U.R.S.S.* **18**, 169.
 Stokes, G. G. 1845 *Trans. Camb. Phil. Soc.* **8**, 287.

Optical diffraction patterns produced by bubble rafts

By J. DYSON

Research Laboratory, Associated Electrical Industries Ltd., Aldermaston, Berks

(Communicated by T. E. Allibone, F.R.S.—Received 13 April 1949)

[Plates 2 to 7]

A method of obtaining diffraction patterns, analogous to X-ray and electron diffraction patterns, from photographs of bubble rafts is described. The results of some early experiments are shown, and the conditions for obtaining patterns of higher resolution and more uniform illumination are deduced. An experimental technique for realizing these conditions is described and the results are shown.

A method is described for viewing the raft in the light of a selected region of the diffraction pattern, and some results of this technique are shown.

INTRODUCTION

The ingenious bubble model illustrating the atomic pattern within crystalline materials which was developed by Bragg & Nye (1947) gives lattice structures so closely analogous to those found in practice and illustrates effects predicted by theory so closely that it was thought that it would be useful to extend the technique to enable diffraction patterns to be obtained from the rafts which would be analogous to the X-ray or electron diffraction patterns produced from the crystalline material itself.

Optical analogues of X-ray or electron diffraction patterns have, of course, been produced by several workers (Pohl 1940; Kathavate 1945; Prins 1931; Sagani 1926), but the production of diffraction patterns by a raft of bubbles is of particular interest because the capillary forces between bubbles are analogous to inter-atomic forces, and also because of the very large number of diffracting centres which may be made to contribute to the pattern.

The bubble rafts are in general only two-dimensional (although Bragg & Nye have experimented with bubble rafts two or three bubbles deep), and hence the diffraction patterns would be more analogous to those given by electron diffraction.

With this in mind, some experiments were performed with the object of obtaining optical diffraction patterns either directly from bubble rafts or from suitably prepared photographs.

EARLY EXPERIMENTS

The first experiment undertaken was that of obtaining diffraction from a perfect hexagonal lattice of diffracting points. Instead of a bubble raft, an assembly of about 1400 $\frac{1}{16}$ in. bearing balls in a glass-fronted frame was used. This arrangement is due to Dr G. A. Geach of this laboratory, and has the advantage that all elements of the lattice are very accurately of the same size, and they can be manoeuvred into the correct lattice arrangement in a way scarcely possible with bubbles. The assembly was photographed on to 16 mm. film at a scale such that the distance between lattice points was about 0.2 mm.

The balls were illuminated by a single photoflood lamp in a reflector, so that each ball was represented by a single image on a very small scale of the light source, formed by reflexion at the convex surface of the ball. The negative was printed on to a 'photomechanical' plate by contact, giving a dark field with a small transparent spot near the centre of each ball image. The contrast was sufficiently high for the transmission of the print to be significant only in the transparent spot. The print was placed in an optical system shown in figure 1. Light from a B.T.H. type ME 250 W. mercury vapour lamp is focused on to a small pinhole and diverges to an achromatic collimator lens. The parallel light traverses the print and enters a telescope lens. The (Fraunhofer) diffraction pattern is formed in the focal plane of this lens and can be examined by an eyepiece. An Ilford 'mercury-green' filter allows of using only the green line of the mercury spectrum, together with a small part of the continuous background, which is rather high in this lamp.

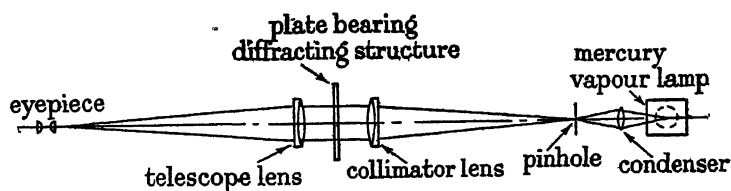


FIGURE 1

The resulting diffraction pattern is shown in figure 2, plate 2, and, of course, consists of a hexagonal array of spots. It can be seen that each spot has faint wings pointing in a radial direction. These are due to the continuous background and are a function of the sharpness of cut-off of the filter used. Generally, though, they are not obtrusive, and in fact are only visible around the brightest spots.

The pattern obtained bears an obvious resemblance to the electron diffraction image obtained when electrons traverse a thin flake of mica. Such a picture, taken by

A. E. Ennos of this laboratory, is shown in figure 3, plate 2. The mica flake used was a number of molecules thick, and so the phenomenon of Bragg reflexion is just beginning to appear; this accounts for the fact that some of the spots are enhanced in intensity. As the print used for the optical diffraction is a true two-dimensional array, no corresponding enhancement of intensity occurs in this case.

If the unfiltered light from the mercury vapour lamp is used, each spot becomes a radial spectrum, of length proportional to its distance from the centre, and the whole forms a beautifully symmetrical pattern of brilliant colours.

An attempt was next made to form diffraction patterns from bubble rafts. To form the pattern direct from the raft itself would be elegant and, for some purposes, desirable, but it is so difficult as to be scarcely practicable. It is difficult to produce very small bubbles of very uniform size, and so the scale of the diffraction pattern would be small, necessitating very perfect optics to attain the required definition. As the light has to traverse the raft, the liquid layer is itself a part of the optical train. The surface through the centres of the bubbles is very far from optically flat, and this would preclude the attainment of anything like the required optical perfection. In addition, the refraction at each bubble, when considered as a negative lens, is so great that the light is spread out over a very large solid angle; hence the amount of light available to form the diffraction pattern is very small.

These considerations seemed to make the chances of success very small. Therefore it was decided to use photographs of the rafts as the diffraction gratings.

A raft was photographed on to 35 mm. film, at a scale suitable to give a bubble-image spacing of 0.2 mm. using two photofloods in reflectors as illumination. Each image consisted of a dark ring enclosing a transparent disk containing two minute dark spots which were the images of the lamps. The negatives were sandwiched between glass plates in Canada balsam. This was necessary as the optical quality of the film surfaces was not nearly good enough to give the required definition.

Results given by this means are shown in figures 4*a* to 7*b*, plates 3 and 4. Figures 4*a*, 5*a*, 6*a*, 7*a* are enlarged prints from the negatives of the bubble raft, and figures 4*b*, 5*b*, 6*b* and 7*b* are the corresponding diffraction patterns.

Figure 4*a* shows a raft photographed immediately after growth. It is seen to consist of a few comparatively large crystallites in a state of considerable internal strain containing a large number of dislocations.

The corresponding diffraction image, figure 4*b*, consists, as might be expected, of a series of spotty rings. The individual spots are comparatively sharply defined and the regions between the rings are dark. The dislocations and strains may be expected to scatter light outside the spots, but the definition is not good enough to show this.

Figures 5*a* and 5*b* show the same raft photographed immediately after being stirred up, with its diffraction pattern. The large crystallites have been broken up and there are areas of amorphous material and a number of small crystallites just beginning to grow.

The diffraction pattern now consists of almost smooth rings which, however, can be seen to consist of a very large number of ill-defined spots. The regions between the rings are now filled with light scattered from the amorphous region. In taking this

exposure an absorbing mask was placed over the centre of the diffraction pattern to reduce the intensity there. Otherwise this was so great that the halation due to the central spot obscured the first ring. As a result the background inside the first ring is darker in the photograph than that between the first and second rings.

Actually, the scattered light fell off monotonically towards the outside of the pattern, as might be expected. Visual inspection of the pattern with an eyepiece gives an impression similar to that shown in the photograph, but this is a subjective contrast effect, caused by the proximity of the overpoweringly bright central spot. This was established by blocking out the central spot when the background intensity appears greater inside the first ring than outside it.

The diffraction pattern may be compared with the corresponding X-ray diffraction patterns for liquids, which exhibit the same general characteristics.

A diffraction figure somewhat similar to figure 5*b* has been given by Pohl (1940), produced by a large number of randomly arranged apertures. The differences between this and figure 5*b* are due to the fact that Pohl's raft had, presumably, no trace of crystal structure.

Figures 6*a*, 7*a* and 6*b*, 7*b* were taken from the same raft at different periods after stirring. The amorphous regions are decreasing, and the effect of this is seen in the diffraction patterns by the clearing of the background. In addition, the crystallites are decreasing in number, and the diffraction rings are becoming more spotty.

A further effect which is very obvious in figure 7*b* is the appearance of six dark spaces in each ring. The positions of these are related to each other in the same way as the spots in figure 2, and evidently indicate that a particular orientation is being avoided by the growing crystallites. This effect can also be seen in an incipient state in figure 6*b*. Although so prominent in the diffraction pattern, this effect is by no means obvious from inspection of the direct photographs (figure 7*a*).

The four white arms radiating from the central spot in figures 4*b*, 5*b*, 6*b* and 7*b* are due to diffraction from the rectangular boundary of the diffracting negative.

DEFECTS IN THE PATTERN

These early experiments give patterns which have two outstanding defects. The resolution of the spots is poor, and there is a great concentration of light towards the centre of the pattern, especially in the central spot. Further experiments were undertaken with the object of overcoming these defects.

BRIGHTNESS AND RESOLUTION

Assume that the diffracting object consists of a square array of N^2 transparent apertures in a black screen, with a linear spacing of d cm. Assume further that if the dimensions of the array be varied the size of the apertures is also varied in such a way as to keep constant the ratio of the areas of the transparent regions to the areas of the opaque regions.

Consider first the condition where the size of a spot formed by diffraction from an area of perfect lattice is determined by the size of the source, i.e. by geometrical optics. This will be the most usual case where a large number of diffracting centres

contribute. The angular diameter of the source is a , and the resolution of the pattern, defined as the ratio of the distance between successive rings of spots to the diameter of a spot, is k . Then

$$k \sim \frac{1}{ad}. \quad (1)$$

The total amount of light collected from the source is proportional to the square of its angular diameter and to the area of the diffracting object:

$$\text{total light} \sim d^2 a^2 N^2. \quad (2)$$

The size of the image on the photographic plate is of importance, because if it is too small the resolution will be spoiled by halation. Let the linear diameter of the first ring be S . Then the brightness of the image will be given by

$$\text{brightness} \sim d^2 a^2 N^2 \frac{1}{S^2}. \quad (3)$$

Substituting from (1) we get $\text{brightness} \sim \frac{N^2}{k^2 S^2}. \quad (4)$

Hence, for a given resolution and image size, the brightness is dependent only on the total number of transparent diffracting areas, if their geometry is constant.

The other case is that where the spot size is fixed by diffraction phenomena, i.e. by the size of the Airy disk. Then, if the source is kept just small enough to ensure this,

$$a \sim \frac{1}{Nd}. \quad (5)$$

The total light collected is given by

$$\text{total light} \sim d^2 a^2 N^2 = \frac{d^2}{N^2} \frac{N^2}{d^2} = \text{constant}. \quad (6)$$

The brightness of the image is thus only dependent on its size. The resolution, however, is given by

$$k \sim \frac{1}{ad} = N, \quad (7)$$

as is well known from the theory of diffraction gratings.

It is thus evident that the important feature is N^2 , the total number of diffracting points.

So far we have not considered the distribution of light among the various rings of the pattern. This is governed by the angular distribution of light from each diffracting element. The diffracting elements in figures 4*a* to 7*a* consist effectively of transparent disks surrounded by thin opaque rings of diameter almost equal to the distance between centres. Such elements give an angular distribution similar to that in an Airy disk. The angular radius of the first dark ring is given by

$$\alpha = \frac{0.61\lambda}{r}, \quad (8)$$

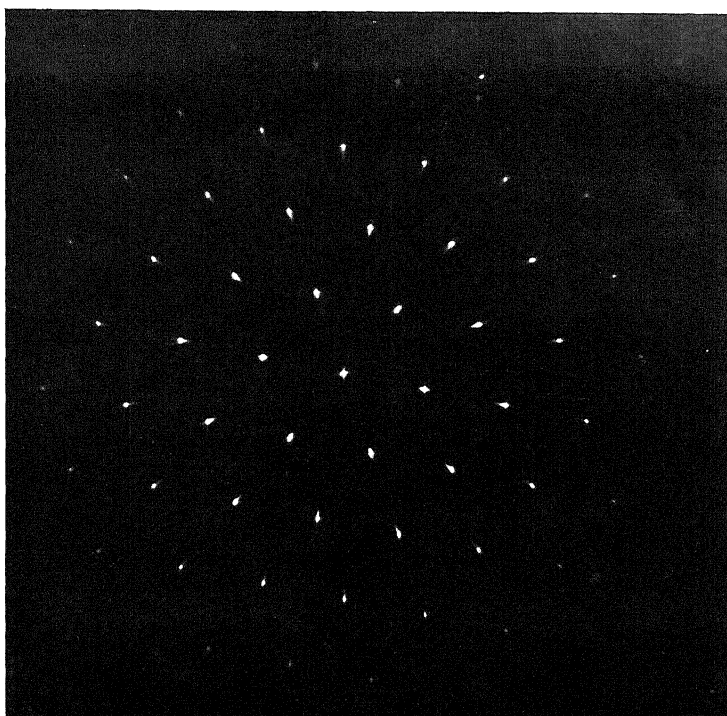


FIGURE 2. Diffraction pattern from a perfect lattice.

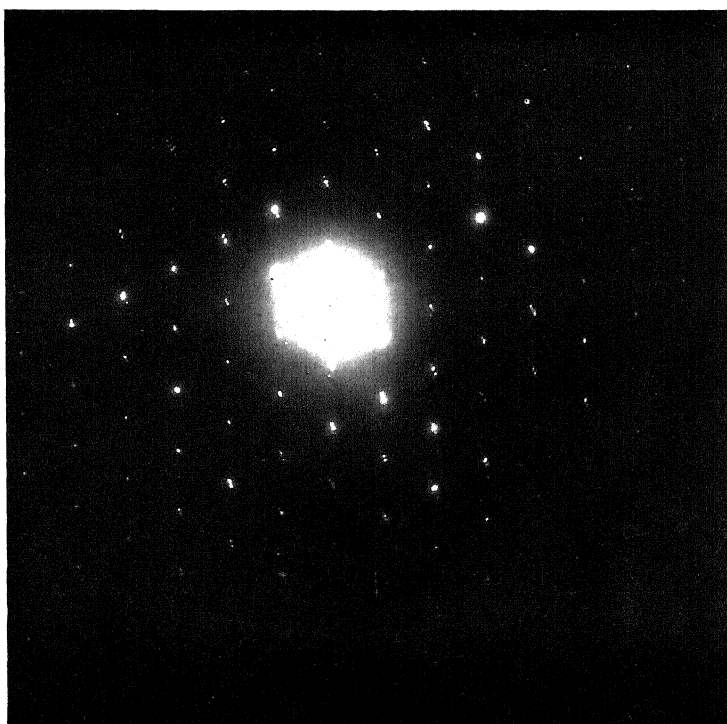
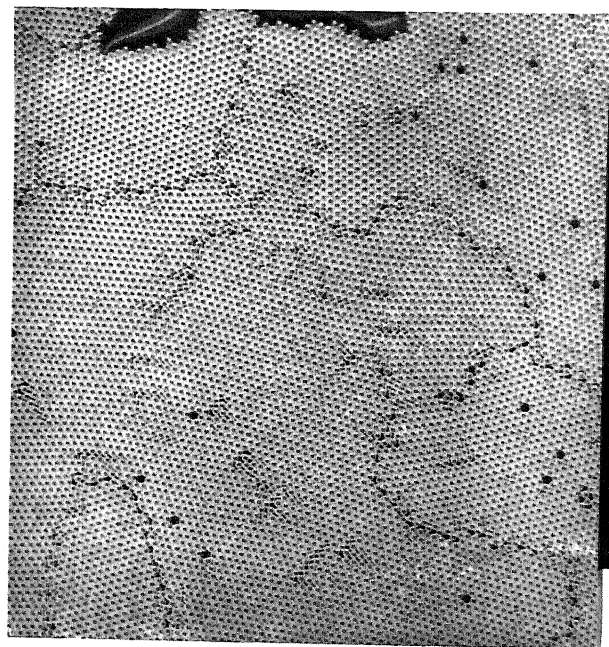
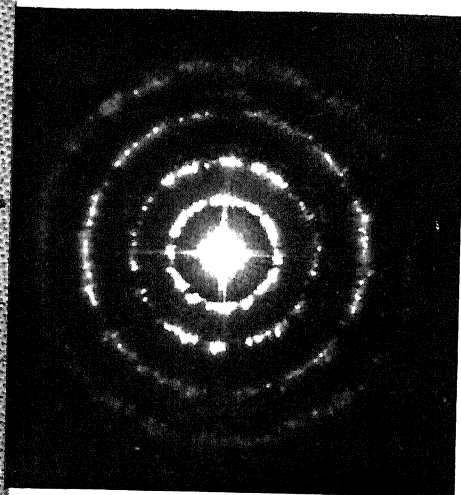


FIGURE 3. Electron diffraction pattern from a mica flake.

(Facing p. 134)

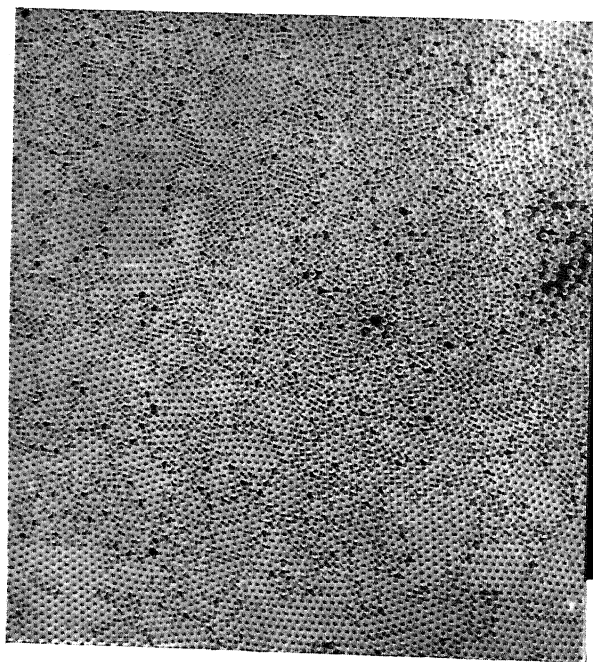


a



b

FIGURE 4. *a*, polycrystalline raft; *b*, diffraction pattern.

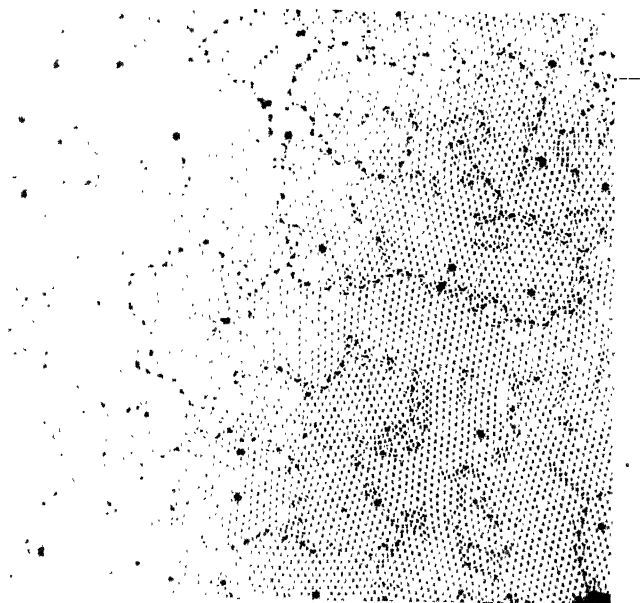


a



b

FIGURE 5. *a*, raft after stirring; *b*, diffraction pattern.

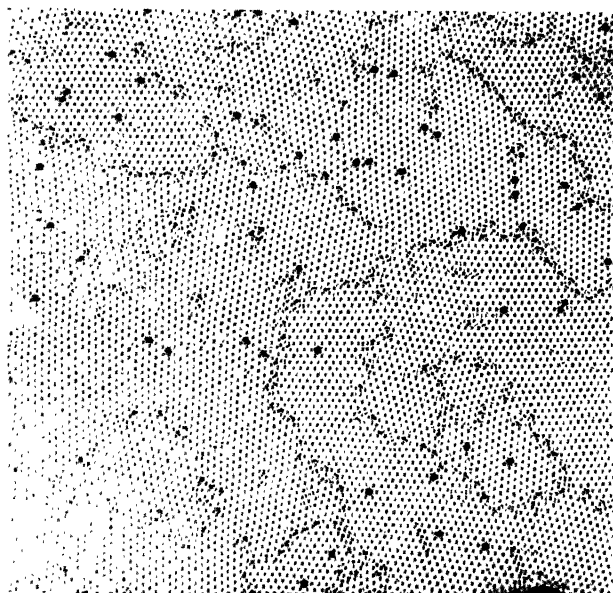


a



b

FIGURE 6. *a*, raft after 'annealing'; *b*, diffraction pattern.



a



b

FIGURE 7. *a*, raft after further 'annealing'; *b*, diffraction pattern.

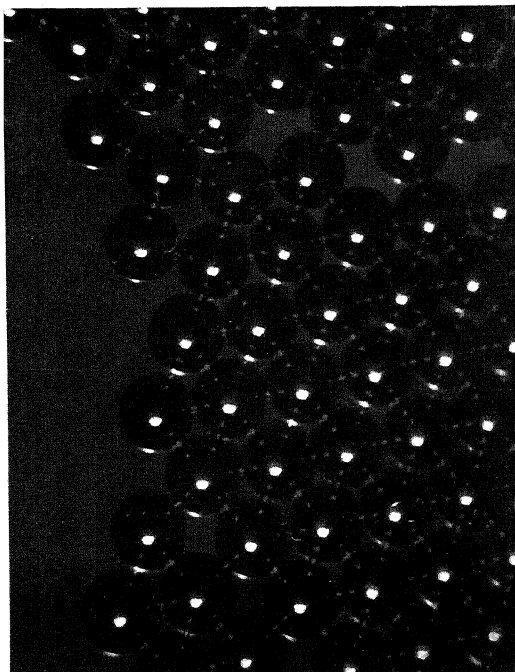


FIGURE 8. Bubbles with back illumination.

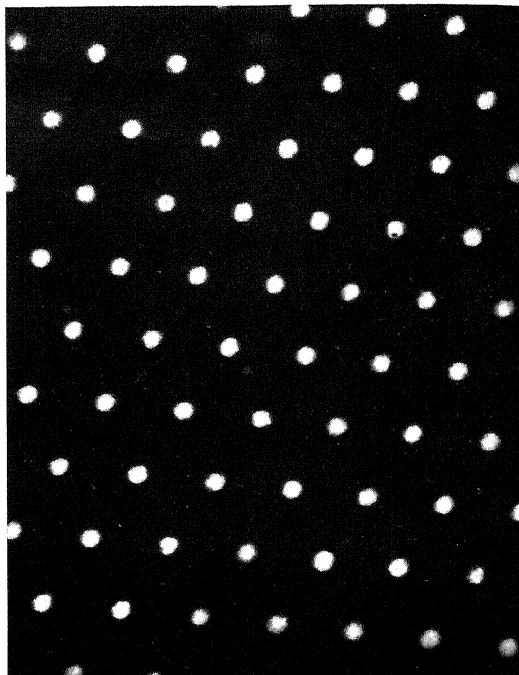
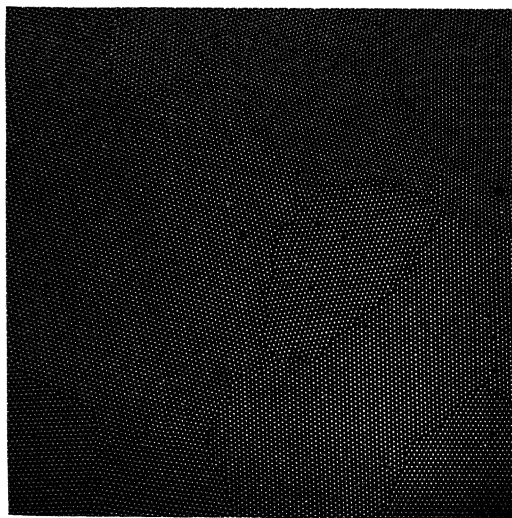
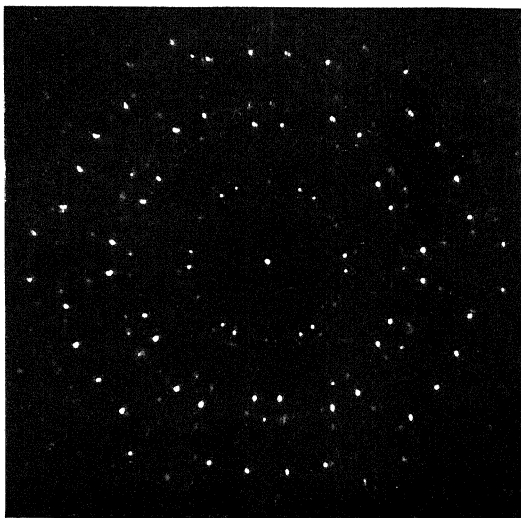


FIGURE 10. Photomicrograph of transparency for diffraction.



a



b

FIGURE 11. *a*, polycrystalline raft with two sizes of bubbles; *b*, diffraction pattern.

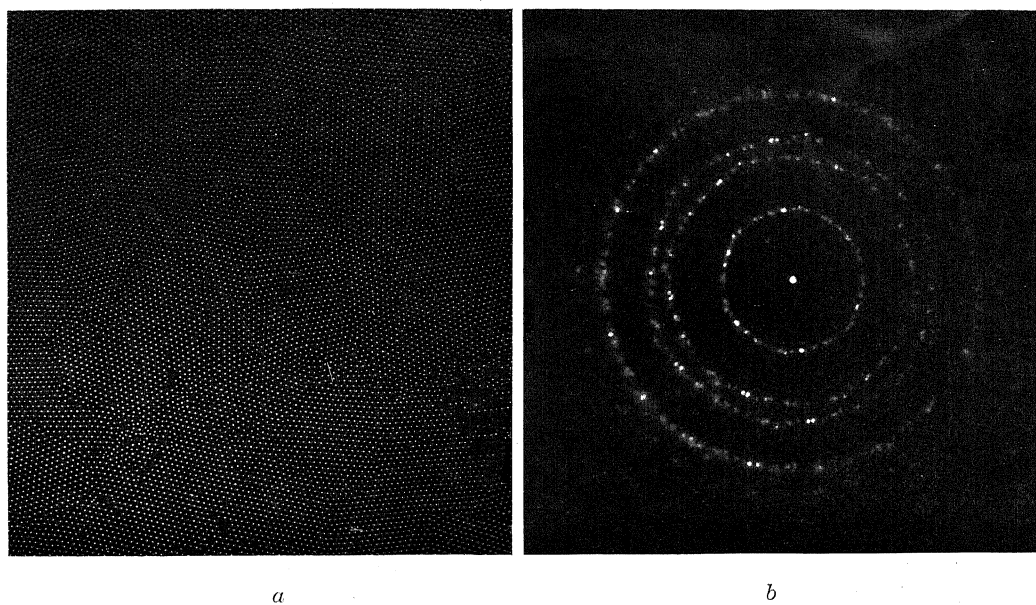


FIGURE 12. *a*, stirred bubble raft; *b*, diffraction pattern.

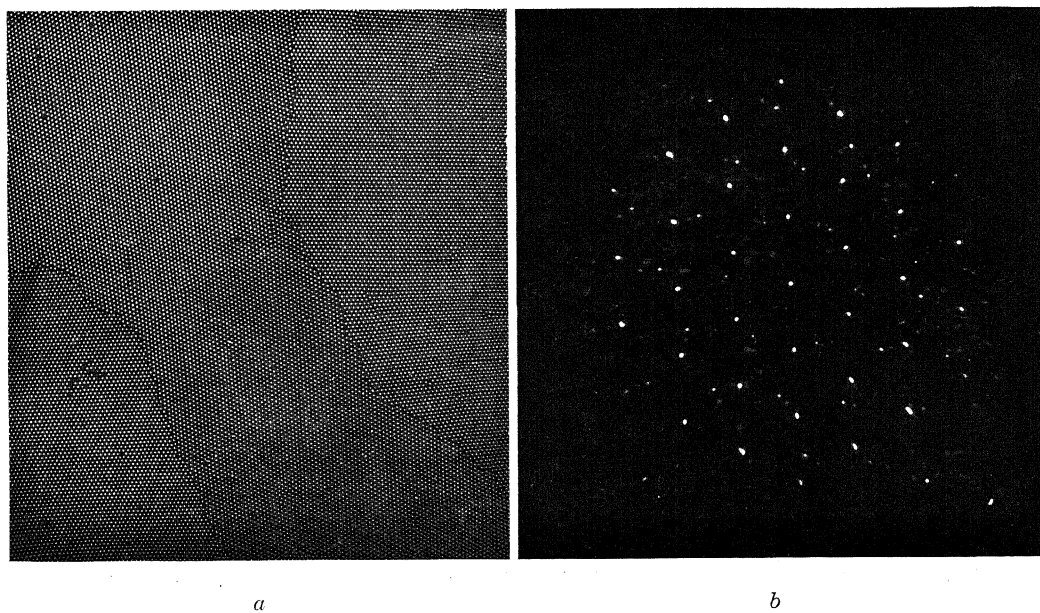


FIGURE 13. *a*, raft of large crystallites; *b*, diffraction pattern.



FIGURE 14. Raft of figure 13*a* viewed through pinhole.

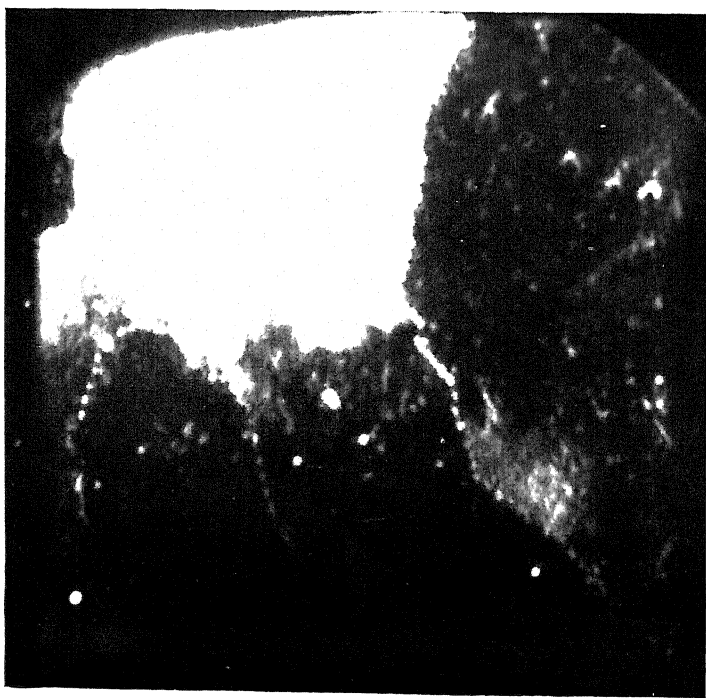


FIGURE 15. As for figure 14 but with different adjustment of pinhole.

where r is the radius of the transparent disk. The angular radius of the first ring in the diffraction pattern is given by

$$\beta = \frac{\lambda}{d}, \quad (9)$$

where d is the distance between centres of the transparent regions.

To obtain sensibly uniform illumination up to, say, the fourth order of interference, we require that the angular radius of the first dark ring of the Airy pattern be larger than the angle corresponding to the fourth order of interference. Thus

$$\frac{0.61\lambda}{r} > \frac{4\lambda}{d},$$

or

$$r < 0.15d. \quad (10)$$

Hence to improve the uniformity of illumination it is necessary to make the transparent spots smaller.

To cut down the intensity of the central spot it is necessary to pay particular attention to the opacity of the background, since the area of the transparent spots is a small fraction of the total area.

It may be noted that, by Babinet's principle, a similar array of small opaque spots on a transparent background will give the same diffraction pattern, but the intensity of the central spot will be very much greater.

EXPERIMENTAL TECHNIQUE

By the above analysis it is seen that what is required to obtain better patterns is to produce bubble rafts with large numbers of bubbles and to photograph these in such a way as to represent each bubble by a small circular transparent spot on an opaque background.

The requirement of a large number of bubbles depends, of course, on the particular phenomenon under investigation. The other requirement is more under control.

If the raft be illuminated by a small light source on the upper side and viewed from the top, a small image of the source is seen by reflexion in the convex surface of each bubble, and another in the concave surface. The predominant feature, however, is a bright ring of nearly the full diameter of the bubble. This is very much brighter than the direct image and cannot be eliminated by adjusting the angle of the light. Hence this method of illumination is of little use. If the bubbles be blown in a transparent trough and a small light source placed underneath, the appearance is as shown in figure 8, plate 5. Each bubble behaves as a negative lens of focal length approximately three times the radius of the bubble and so forms a small erect image of the light source. Further images are formed by multiple reflexion at the convex surfaces of adjacent bubbles, but these are so small that they do not cause much trouble. The light source is displaced to one side away from the axis. The three-cornered areas between the bubbles then appear dark, while the direct image is still visible. It is shifted to one side, but this occurs in every bubble and so is equivalent to a slight lateral shift of the pattern.

To make use of this arrangement the bubble raft is blown in a Perspex trough about 15×10 cm. and about 1 cm. deep. For convenience, two front-reflexion

mirrors are mounted at 45° to the vertical, one above and one below the trough. The light source is a photoflood about 3 ft. from the trough, and the raft is photographed by means of a camera at a distance such as to give a spacing of about 0.2 mm. between bubble images. The experimental arrangement is shown in figure 9.

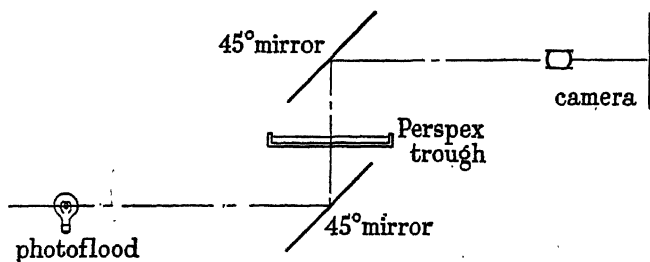


FIGURE 9

The camera lens is thrown out of focus just sufficiently to cause each image of the light source to appear as a small round disk of the size required by equation (10). For this to be successful, the lens must be well corrected; a Goerz Doppel Anastigmat working at $f/8$ has been found to be satisfactory.

The plate used was Ilford Special Rapid Panchromatic, developed in I.D. 8 (maximum contrast) developer. A positive contact print is made on Ilford thin-film half-tone stripping plate, also developed in I.D. 8. The overall gamma of the double process is of the order of 10 and is sufficiently high to suppress all detail except the circular out-of-focus disks. The result is a pattern of small circular transparent disks on an opaque background, as is shown in the photomicrograph (figure 10, plate 5). The fact that the disks have slightly soft edges is of no importance, as the effect of this is only to suppress the outer rings of the Airy pattern, leaving the central lobe almost unchanged.

The glass of a photographic plate is not of sufficiently high quality to allow of sharp patterns being formed, so the emulsion is peeled off the stripping plates, having been cut round the required region with a knife. The emulsion is immersed in a bath of methylated spirit which has been diluted with about an equal volume of water. This enables the gelatine to take up enough water to adhere to a glass surface and yet to retain its strength. An optically flat glass plate (selected pieces of $\frac{1}{8}$ in. plate glass can be used) is slid under the emulsion and lifted out. Wrinkles are gently pressed out with the fingers and the plate stood on edge to dry.

When dry, a drop of Canada balsam is placed in the centre of the emulsion and another glass flat pressed down over the top. A small weight is placed thereon and the whole placed in an oven at about 50°C overnight. The balsam will spread over the whole area; the excess is wiped off at the edges and, after cooling, the resulting sandwich is rigid enough if handled carefully.

EXPERIMENTAL RESULTS

A bubble raft photographed by this technique and its diffraction pattern are shown in figures 11*a* and 11*b*, plate 5. This raft consisted largely of areas of almost perfect lattice, and this is reflected in the rings of sharp spots shown in the diffraction pattern.

It will be noted that all the rings are doubled; this is because two different sizes of bubbles are present, as can be seen from close inspection of figure 11*a*.

Figures 12*a* and 12*b*, plate 6, show a raft taken just after stirring, together with its diffraction pattern. The rings of diffuse spots as shown in figure 5*b* appear once more, but the background shown in figure 5*b* does not appear. The reason for this seems to be that a small delay occurred between stirring and taking the photograph in the case of figure 12*a*, and the amorphous regions had time to crystallize into a large number of small crystallites.

Figures 13*a* and 13*b*, plate 6, are of a raft consisting of only three large crystallites and a small one. The diffraction pattern consists of only a comparatively small number of spots, but these are seen to be broken and surrounded by a diffuse nebulosity. This effect is evidently due to lack of perfection in the crystallites, and the question as to what form of imperfection gave rise to this phenomenon suggested the investigation described in the next section.

DIRECT OBSERVATION OF IMPERFECTIONS

An imperfect lattice such as that of figure 13*a* constitutes an almost periodic structure, from which a Fourier analysis gives sets of terms grouped closely around the discrete values which would represent the corresponding perfect lattice. One such set of terms gives information, generally speaking, about the broad features of the lattice, whereas the overall distribution of sets of terms contains information as to the microscopical features, such as the size and shape of the dots (see the foregoing analysis giving the relationship between the spot size and the distribution of light between the various rings).

Thus, one particular spot in the pattern, together with its associated nebulosity, can be expected to yield information about the broad features, such as shape and imperfections, of one crystallite.

The technique for applying this principle consists of mounting a thin opaque plate containing a pinhole about 0.5 mm. diameter in the focal plane of the telescope lens. The plate is mounted on centring screws so that it can be moved about to coincide with any one of the spots of the diffraction pattern. Then, with the eyepiece removed, the eye is placed close to the pinhole, when the image of the bubble raft is seen in the aperture of the telescope lens, illuminated by only the light which forms the particular spot selected.

The resulting phenomena are easier to interpret if the unfiltered light is used. Under these conditions it is possible, for certain positions of the pinhole, for the whole of the raft to be illuminated at once by light of different wave-lengths and in different orders of diffraction. Detail is then seen both by intensity and colour contrast. The resulting picture resembles the coloured counties on a map.

This may be illustrated with reference to figure 14, plate 7, which shows the raft of figure 13*a* viewed under these conditions. The individual bubble images cannot be distinguished, as only one diffracted beam from each area of lattice is used. Thus the experiment is a corroboration of Abbe's theory of the microscope. The grain boundaries, however, show up clearly as rows of approximately equally spaced dots.

By making a transparency of figure 14, to the same scale as figure 13*a* and superimposing them these dots can be identified as those regions along the grain boundaries which are nearly, but not quite, large enough for the inclusion of an extra bubble.

In figure 13*a* two dislocations can be seen across the waist of the central crystallite. These cause the top and bottom of the crystallite to be rotated slightly with respect to each other. This is shown clearly in figure 14, the upper region being brighter than the lower. To the naked eye, the upper region was yellow and the lower one red, the resulting contrast being very clear. The boundary between the two regions is seen to be quite sharp, in fact as sharp as the limited resolution of the picture will allow. The two dislocations show up as dark streaks. Another boundary is shown at the upper left of this crystallite where an 'arm' branches off to the left.

The right-hand crystallite is complex in structure. There is a line of dislocations across the upper portion shown very clearly in figure 14 as a row of 'puckers'. The boundary between the two regions split off by the dislocations is not sharp in this case. Visually, the upper region was dark blue and the lower region light blue. Towards the bottom of this crystallite is a complex structure of dislocations. The resulting pattern of boundaries was almost impossible to photograph in monochrome, but showed up visually, as a complicated pattern of green, red and light blue areas. The dislocations, however, can be seen in figure 14 as a number of dark markings. These give a 'plastic' appearance to the raft as if it consisted of a puckered and torn sheet of fabric. All the dislocations may be seen by close inspection of figure 13*a* but the boundaries between areas of perfect lattice cannot be followed so clearly.

Figure 15, plate 7, gives another view of the same raft with a different pinhole setting. Only the upper region of the central crystallite is brightly illuminated; the boundary does not show so clearly because of scattered light, as the upper region was over-exposed to bring out details elsewhere. The grain boundaries now show up in rows of bright dots. It should be noted that these dots represent structures of the dimensions of one bubble site, and the resolution of the picture is not sufficient to show individual bubbles; yet, as these dots indicate deviations from perfect regularity, they show up vividly both on a light and on a dark ground.

The dislocations in the right-hand crystallite now show up by themselves as V-shaped markings. Visually they appear as bright golden yellow marks on a 'midnight blue' background, their contrast being much higher than the photograph suggests.

As the centring screws are operated the colours change in such a manner that the colour change is almost simultaneous over each area of perfect lattice. The effect is that of tilting a solid object with plane facets, each facet catching the light in turn. Crystal boundaries appear as sharp edges, whereas the boundaries between areas in the same crystallite appear rounded to a greater or less degree. A few minutes of such observation give a much better idea of the nature of a raft than can be obtained from simple inspection, or than can easily be described in words.

CONCLUSION

The experiments here described have two potential uses. In the first place, it may be possible, from a study of a sufficient number of rafts and their associated diffraction patterns, to obtain some insight into the relationships between observed phenomena in X-ray and electron diffraction and the crystal imperfections which cause them.

In the second place the method of inspection of the raft by means of the light contained in a restricted portion of diffraction pattern offers a ready method of investigating the effects produced by dislocations on the surrounding lattice and leads to a clearer understanding of the way in which complex diffraction spots are formed.

The value of this method lies in its directness; the imperfection producing a given phenomenon in the diffraction pattern can be seen clearly without ambiguity.

A limitation to the usefulness of the analogy is imposed by the fact that only two-dimensional lattices can be treated in this way. However, this can be quite a close approximation to conditions obtaining in electron diffraction. It was not thought worth while to try to extend the method to three-dimensional lattices, partly on account of the very great experimental difficulties, but also because it would be almost impossible to form a clear idea of the imperfections existing in a solid block of lattice.

The author wishes to thank Dr T. E. Allibone, F.R.S., for permission to publish this paper.

REFERENCES

- Bragg, L. & Nye, J. F. 1947 *Proc. Roy. Soc. A*, **190**, 474.
Pohl, R. W. 1940 *Optik*, p. 100, figure 232.
Kathavate, Y. V. 1945 *Proc. Indian Acad. Sci. A*, **21**, 253.
Prins, J. A. 1931 *Naturwissenschaften*, **19**, 435.
Sagani, C. M. 1926 *Phil. Mag.* **1**, 321.

BAKERIAN LECTURE
A region of biosynthesis

BY H. RAISTRICK, F.R.S.

London School of Hygiene and Tropical Medicine

(Delivered 12 May 1949—Received 31 May 1949)

[PLATES 8 TO 11]

INTRODUCTION

The particular region of biosynthesis about which I propose to speak is that part of the kingdom of living things occupied by the fungi, particularly those fungi popularly known as 'moulds', with occasional reference to some higher fungi and also to lichens, which are of course symbionts of algae and fungi. It is, I think, not inappropriate that this should form the subject of a Bakerian Lecture, since the founder of this lecture, Mr Henry Baker, F.R.S., who was an ardent microscopist, described in 1742 the spores of the fungus *Lycoperdon*,* the common puff ball⁽¹⁾.† In the following year he gave a good description of *Pilobolus*, a mould of the mucor type, which he found growing on a culture of black mud from the river Thames⁽²⁾.

The exact position of fungi in the scheme of living things is still doubtful, since, to quote Dr John Ramsbottom⁽³⁾, 'if organisms must be either plants or animals, then fungi are plants with a nutrition resembling that of animals as they do not possess chlorophyll. If, however, chlorophyll is the hall mark of plant phylogeny the fact has to be faced that fungi probably never possessed it.' Because of the absence of chlorophyll, fungi can be cultivated only on media containing pre-formed organic matter, but this very fact makes moulds particularly suitable for biochemical investigation, since they grow well on very simple media.

In spite of this fact, real interest in the chemistry of moulds may be said to date from 1891 with Carl Wehmer's classical observations that when *Aspergillus niger* is grown on sugar solutions oxalic acid is formed in considerable quantities^(4, 5, 6, 7), and that citric acid is a metabolic product of certain mould species to which he gave the generic name *Citromyces*^(8, 9). For the next thirty years only an occasional paper appeared on this subject, and this was the position when I was fortunate enough to be able to begin systematic work on the subject in 1923 with a small team of enthusiastic colleagues in the research laboratories of Nobel's Explosives Company Ltd., Ardeer, Scotland. This work was carried on there until 1929⁽¹⁰⁾, since when it has been continued in the Department of Biochemistry at the London School of Hygiene and Tropical Medicine.

* A list of species cited in the text with the authorities for each is printed in the appendix to the lecture.

† Because of the large number of literature references cited, the custom usually followed in this journal of quoting in the text the names of authors and year of publication has been altered. Each publication referred to is given a number and the same numbering is used in the list of references at the end of the lecture.

EXPERIMENTAL METHODS

The general plan of the work was the investigation of the products of metabolism of pure cultures of species and strains of moulds when grown under controlled conditions on simple, reproducible and chemically defined culture media. We have used almost exclusively one or other of the two following media:

Czapek-Dox medium		Raulin-Thom medium	
glucose	50 g.	glucose	75 g.
NaNO ₃	2 g.	tartaric acid	4 g.
KH ₂ PO ₄	1 g.	ammonium tartrate	4 g.
KCl	0.5 g.	(NH ₄) ₂ HPO ₄	0.6 g.
MgSO ₄ ·7H ₂ O	0.5 g.	(NH ₄) ₂ SO ₄	0.25 g.
FeSO ₄ ·7H ₂ O	0.01 g.	K ₂ CO ₃	0.6 g.
distilled water	1 l.	MgCO ₃	0.4 g.
		FeSO ₄ ·7H ₂ O	0.07 g.
		ZnSO ₄ ·7H ₂ O	0.07 g.
		distilled water	1.5 l.

It will be seen that in the Czapek-Dox medium the only source of organic matter is glucose, and in the Raulin-Thom medium glucose + tartaric acid. The other elements essential for growth, nitrogen, phosphorus, potassium, sulphur, magnesium and the trace elements are present as mineral salts. Hence, any of the organic metabolites which I shall describe must have arisen by synthesis from glucose or glucose and tartaric acid.

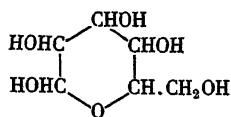
The medium, distributed in 350 ml. amounts in 1 l. conical flasks plugged with cotton-wool, is sterilized, inoculated with a pure culture of the species or strain of mould under investigation, and incubated in the dark at 24° C. Hence all the cultural conditions are standardized and reproducible, the only variant being the species or strain of mould used, so that I think it may be claimed fairly that structural differences in chemically related metabolic products may be attributed to differences in the enzyme systems present in different species or in strains of the same species.

The cultures are harvested when the residual glucose is about 0.5 %. The mould mycelium is separated by filtration from the culture fluid without sterilization, washed with water, pressed and dried in a vacuum oven at 40° C. It is then ground to a fine powder, extracted with suitable solvents, and any extractives are purified by the usual chemical methods.

Any metabolites in the clear culture filtrate and mycelium washings are isolated by extraction with suitable solvents, or by the addition of suitable precipitants, or even by vacuum evaporation when they may occasionally crystallize out.

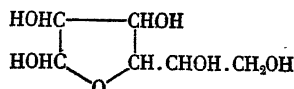
During the last twenty-six years a large number, certainly approaching two hundred, of mould metabolic products have been isolated in a state of purity by my colleagues and myself. It will not be possible to deal with all of them to-day, nor would I wish to weary you with a mere catalogue of fungal metabolic products. My purpose is rather to try to trace the interrelationships in chemical structure between the different types of mould metabolic products described by ourselves and other workers, and to suggest how some of the more complex ones may arise by bio-synthesis from the simpler ones.

KOJIC ACID, PATULIN, 2-HYDROXYMETHYL-5-FURANE CARBOXYLIC ACID AND TERREIN



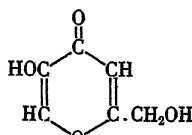
I

glucopyranose



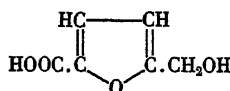
II

glucofuranose

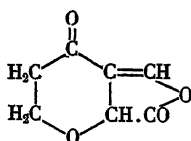


III

kojic acid

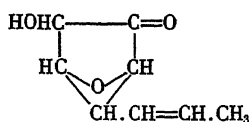


IV

2-hydroxymethyl-5-furane
carboxylic acid

V

patulin (etc.)



VI

terrein

The first group of mould metabolic products is chosen because its members show some of the closest relationships occurring in mould metabolites to the glucose molecule, either in its pyranose (structure I) or furanose form (II). Their inter-relationships in structure are not particularly clear, though they all contain a bridged oxygen ring.

Kojic acid (III), originally isolated by Yabuta from culture filtrates of *Aspergillus oryzae* (11), has been described since then by many workers as a metabolite of a large number of different species of *Aspergillus* and is obtained occasionally in yields of over 50 %.

Patulin (V) was discovered independently and almost simultaneously during the last war in a number of laboratories in different countries. This is the reason why it now appears in the literature under a variety of names—clavacin from *A. clavatus* (12), clavatin from *A. clavatus* (13), claviformin from *Penicillium claviforme* (14, 15), expansin from *P. expansum* (16, 17) and patulin from *P. patulum* (18). Kojic acid and patulin may, I think, be legitimately regarded as derivatives of glucopyranose (I).

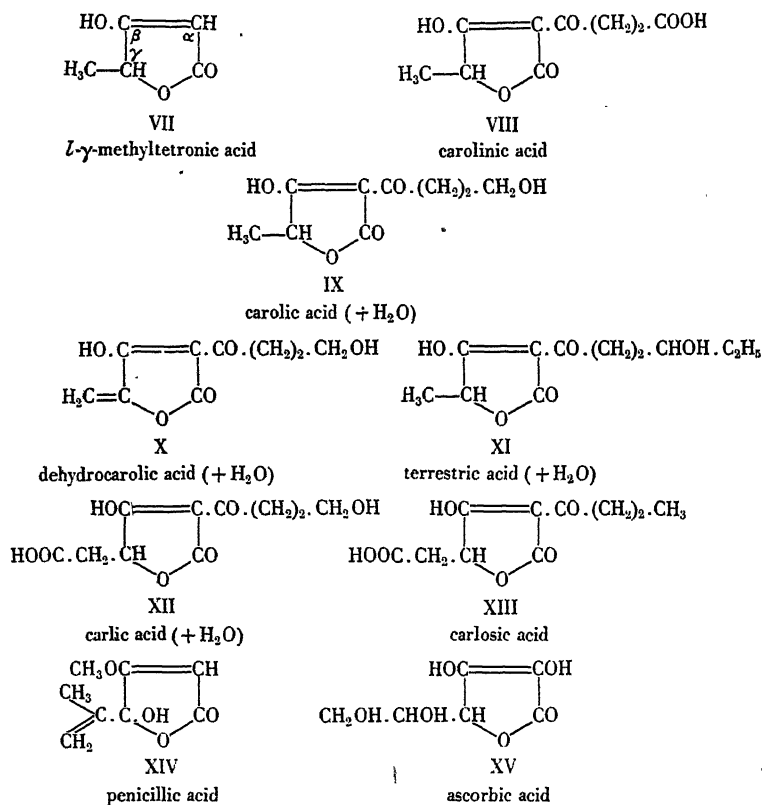
2-Hydroxymethyl-5-furane carboxylic acid (IV) obtained from *Aspergillus glaucus*, *A. clavatus*, *A. niger*, *A. oryzae* and *A. wentii* by Sumiki (19, 20), and terrein (VI) from *A. terreus* (21, 22) may be regarded, though perhaps with less certainty, as derivatives of glucofuranose (II).

DERIVATIVES OF TETRONIC ACID

The next group of metabolic products is a series of nine derivatives of tetronic acid, all of which were isolated from the culture filtrates of species of *Penicillium* or *Aspergillus*. The acids from *Penicillium charlesii*, *l*- γ -methyltetronic acid (VII), carolic acid (IX), carolinic acid (VIII), carlic acid (XII) and carlosic acid (XIII) were obtained in a total yield of about 14 % of the glucose metabolized (23). The

structural formulae for a number of these tetronic acids are given in their hydrated form for ease of comparison.

We regard γ -methyltetronic acid (VII) (24) as the parent substance of this group. Carolic acid (IX) (25) may be regarded as derived from it by substitution of the 4-carbon side-chain in the α position. Carolinic acid (VIII) (25) differs from carolic acid only in having a carboxyl group in place of a carbinol group, and dehydrocarolic acid (X) from *P. cinerascens* (26) in having a γ -methylene group in place of the γ -methyl group. Terrestric acid (XI) from *P. terrestris* (27) is clearly an ethyl carolic acid. Carlic acid (XII) (28) has the same structure as carolic acid, with the exception that a carboxymethyl group takes the place of the γ -methyl group in carolic acid. Carlosic acid (XIII) (28) may be regarded as a reduced carlic acid having a terminal methyl group in place of CH_2OH . Carlosic acid was also obtained along with dehydrocarolic acid from *P. cinerascens* (26).



Penicillic acid (XIV) was described originally in 1913 by the American workers Alsberg & Black (29) who obtained it from *P. puberulum*. It was isolated from *P. cyclopium*, and its molecular structure was determined in my department (30). This structure has since been confirmed by an elegant synthesis carried out by Dr R. A. Raphael in 1948 (31) in Messrs May and Baker's research laboratories and at the Imperial College, London. The presence of a methoxy group will be noticed, a feature of common occurrence in mould metabolic products. Its structural relationship to γ -methyltetronic acid is obvious.

Ascorbic acid (XV) was identified along with much larger amounts of citric acid by Geiger-Huber & Galli in 1945 from cultures of *Aspergillus niger* (32). It is surely more than a coincidence that these two substances are also present together in the juice of citrus fruits.

TABLE 1. ACID HYDROLYSIS PRODUCTS OF TETRONIC ACIDS

acid	mols.	mols.	mols.
γ -methyltetronic	1 CO ₂	1 acetoin	—
carolic	1 CO ₂	1 acetoin	1 butyrolactone
carlic	2 CO ₂	1 acetoin	1 butyrolactone
terrestrial	1 CO ₂	1 acetoin	1 <i>l-n</i> -hexanolactone (ethyl-butyrolactone)
carolinic	1 CO ₂	1 acetoin	1 succinic acid
carlosic	2 CO ₂	1 acetoin	<i>n</i> -butyric acid
dehydrocarolic	1 CO ₂	1 diacetyl	1 butyrolactone

An outstanding feature of most of the members of this group of substances is the ease with which the whole molecule splits apart on hydrolysis with boiling, dilute mineral acids. The hydrolysis products of seven of them summarized in table 1 are produced almost quantitatively, except with dehydrocarolic acid. All of them give either one or two molecules of carbon dioxide. All of them give one molecule of acetoin with the exception of dehydrocarolic acid, from which diacetyl is produced. The distinctive features of the different molecules are seen in the hydrolysis products in the last column—butyrolactone, *l-n*-hexanolactone, succinic acid and *n*-butyric acid. This, together with much supporting evidence, justifies we believe the structural formulae assigned to the different tetronic acids. So far none of the seven given in table 1 has been synthesized, except γ -methyltetronic acid.

DERIVATIVES OF CITRIC ACID

We now turn to citric acid (XVI) and a number of other fungal metabolic products which are structurally related to it. Citric acid itself has been known to be a mould metabolite for over half a century, and has been reported by numerous workers from a large number of different species in different genera of moulds and from some higher fungi. Several thousand tons of it are now manufactured annually in yields of the order of 90 % of the theoretical by the fermentation of sugar solutions with *A. niger*.

Itaconic acid (XVII) was first isolated by Kinoshita from cultures of *A. itaconicus* in 1931 (33, 34) and in 1939 in my department from *A. terreus* (35). Since then, American workers using different strains of *A. terreus* and modified cultural conditions have reported yields of itaconic acid of the order of 30 % of the theoretical (36).

Spiculisporic acid (XVIII) has been obtained from three different species of *Penicillium*, *P. spiculisporum* (37), *P. crateriforme* (38) and *P. minio-luteum* (39). It may be regarded as a decyl-homocitric acid (40), as is indicated, *inter alia*, by the fact that γ -ketopentadecic acid (XIX) is also a metabolite of *P. spiculisporum* (37) and is formed from spiculisporic acid by oxidation with potassium permanganate.

Minioluteic acid (XX) occurs along with spiculisporic acid in the culture filtrate of *P. minio-luteum* (39). It is a decyl- α -hydroxycitric acid.

Caperatic acid is a monomethyl ester of nor-caperatic acid (XXI), but it is not known at present which of the three carboxyl groups is esterified in caperatic acid itself. It was described as a constituent of the lichen *Parmelia caperata* by Hesse in 1898 (41), and the structure of nor-caperatic acid, which is clearly a tetradecyl-citric acid, was established by Asano & Ohta (42, 43).

Agaricic acid (XXII), from the higher fungus *Fomes officinalis* (*Polyporus officinalis*), has been known for a century. It is clearly a hexadecylcitric acid (44).

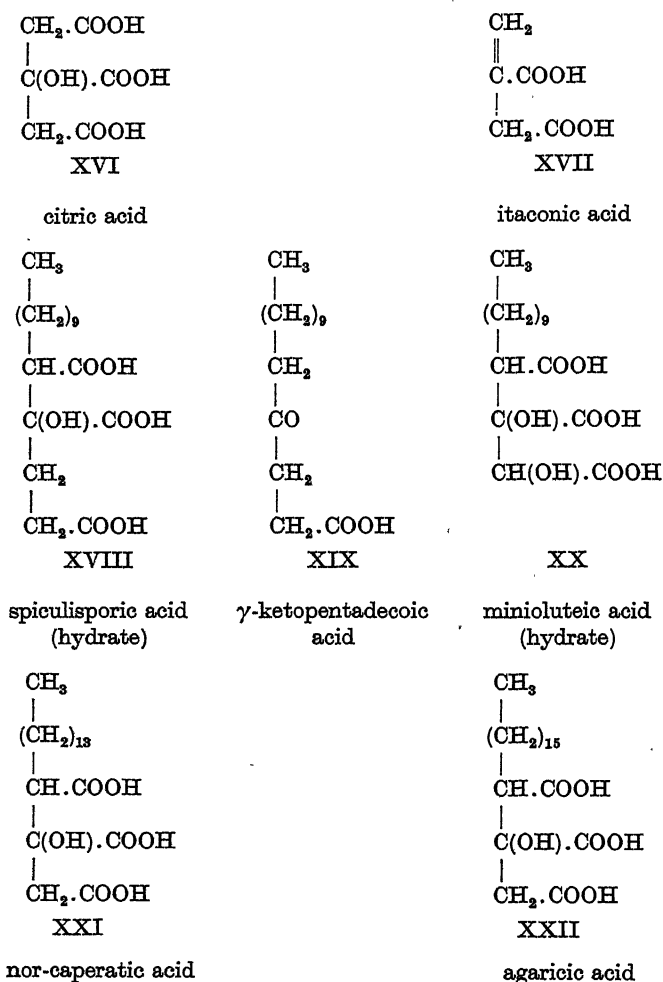
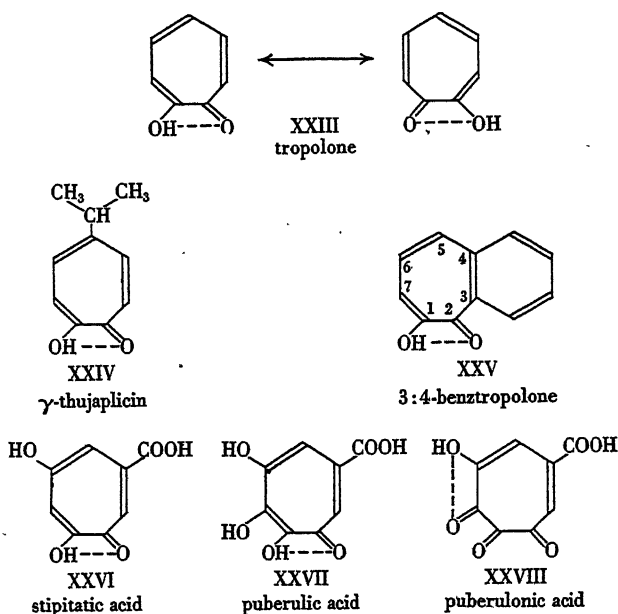


TABLE 2. SUBSTITUTED SUCCINIC ACIDS AS BREAKDOWN PRODUCTS OF MOULD METABOLIC PRODUCTS OR LICHEN ACIDS

mould metabolic product or lichen acid	substituted succinic acid
terrein	<i>d-n</i> -propyl
palitantin	<i>d-n</i> -heptyl
minioluteic acid	<i>n</i> -decyl
nor-caperatic acid	α -methyl- α' - <i>n</i> -tetradecyl
agaricic acid	α -methyl- α' - <i>n</i> -hexadecyl

Before leaving the aliphatic mould metabolites there is a matter of some structural significance to which I wish to draw your attention. In the course of the investigation of their molecular constitution it has been found that a number of products from moulds, higher fungi and lichens give rise to substituted succinic acids on degradation. These are summarized in table 2. Thus tetrahydroterrein gives *n*-propylsuccinic acid (22). Palitantin, $C_{14}H_{22}O_4$, a metabolite of *Penicillium palitans*, is a complex unsaturated dihydroxyaldehyde of at present unsettled molecular constitution (45). Its reduction product, tetrahydropalitantin, yields *n*-heptylsuccinic acid on oxidation. The long side-chains in the substituted succinic acids obtained by the breakdown of minioluteic (39), nor-caperatic (43) and agaricic (43) acids contain respectively 10, 14 and 16 carbon atoms.

DERIVATIVES OF TROPOLONE

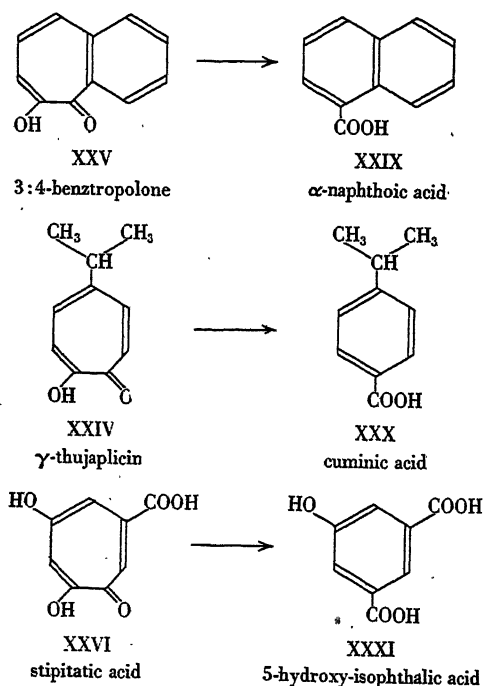


We now come to a group of three quasi-benzenoid mould metabolites, stipitatic acid, $C_8H_6O_5$, from *Penicillium stipitatum* (46), and puberulic acid, $C_8H_6O_6$, and puberulonic acid, $C_8H_4O_6$, from *P. puberulum* (47). A large amount of work was carried out by my colleagues, including the late Professor George Barger (48), and myself without our being able to offer a plausible structural formula for any of them. Thanks, however, to a novel suggestion made in 1945 by Dr M. J. S. Dewar of Oxford it now appears probable that they all contain a 7 carbon ring and are derivatives of what Dr Dewar suggests should be called tropolone (XXIII) (49, 50).

Since 1945 evidence has accumulated supporting Dr Dewar's suggestion. Colchicine is believed to contain a tropolone ring (50, 51, 52). Erdtman & Gripenberg of Stockholm have proved conclusively that α -, β - and γ -thujaplicins (XXIV) from *Thuja plicata*, the western red cedar tree, are respectively α -, β - and γ -isopropyl-

tropolones (53, 54). A few weeks ago Professor J. W. Cook of Glasgow announced the synthesis of 3:4-benzotropolone (XXV) (55).

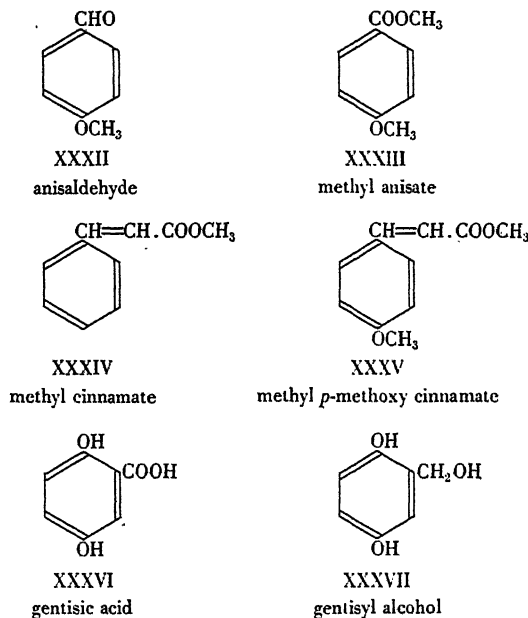
The chemical properties of stipitatic, puberulic and puberulonic acids agree with their formulation as tropolone derivatives. Dr Aulin-Erdtman of Stockholm (private communication) has determined their ultra-violet absorption spectra and concludes from the similarity in spectral type between them and the three thujaplicins that they may all be put into one and the same structural type. The structure assigned to stipitatic acid (XXVI) may, I think, be regarded as a fairly certain one. That suggested for puberulic acid (XXVII) is purely speculative, since there is at present no clear evidence as to the positions occupied by the carboxyl and the two hydroxy groups. Accepting this formulation, however, puberulonic acid may well have the structure XXVIII, since Barger obtained evidence, which unfortunately was not published and is now lost, that puberulonic acid is a quinonoid body closely related to puberulic acid.



A striking characteristic of the tropolones is the ease with which they are almost quantitatively converted into the corresponding substituted benzoic acids on heating with very strong aqueous potash. Thus 3:4-benzotropolone (XXV) yields α-naphthoic acid (XXIX) (55), γ-thujaplicin (XXIV) gives cuminic acid (XXX, *p*-isopropyl-benzoic acid) (54), and stipitatic acid (XXVI) gives 5-hydroxy-*iso*-phthalic acid (XXXI) (46). While it is not suggested that this mechanism explains the formation of true benzenoid compounds by moulds, we may regard stipitatic acid as one possible bridge between them and the aliphatic mould metabolites. So let us now examine some of the benzenoid mould metabolites.

SIMPLE BENZENE COMPOUNDS WITH ONE CARBON CHAIN

My colleague, Dr J. H. Birkinshaw, and Dr W. P. K. Findlay, of the Forest Products Research Laboratory, Princes Risborough, have recently been engaged in a study of the biochemistry of the wood-rotting fungi. They have isolated from some of these higher fungi a number of interesting metabolic products. In particular, they have shown that laboratory cultures of *Trametes suaveolens* and *Lentinus lepideus*, both of which have characteristic aromatic odours, contain the series of four simple benzenoid compounds, XXXII, XXXIII, XXXIV, XXXV, which are clearly closely related to each other structurally. Cultures of *Trametes suaveolens* owe their odour to anisaldehyde (XXXII) and methyl anisate (XXXIII)⁽⁵⁶⁾, and those of *Lentinus lepideus* to the methyl esters of cinnamic acid (XXXIV) and *p*-methoxycinnamic acid (XXXV)⁽⁵⁷⁾. It will be seen that all four metabolites contain either a methyl ether or a methyl ester grouping or both.

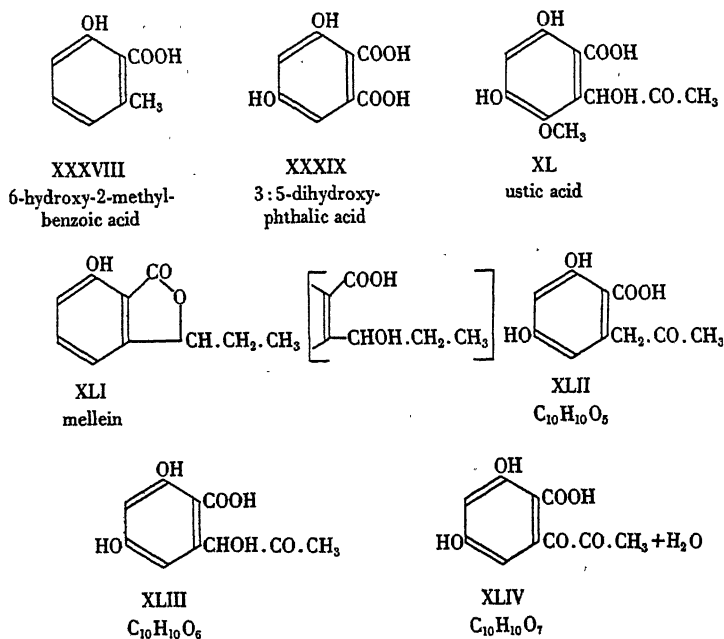


Two other simple benzene compounds with one carbon chain—gentisic acid (XXXVI) and gentisyl alcohol (XXXVII)—are metabolites respectively of *Penicillium griseo-fulvum*⁽⁵⁸⁾ and *P. patulum*⁽⁵⁹⁾. It is a curious fact that such a simple substance as gentisyl alcohol has not, so far as we can ascertain, been described previously, although the corresponding acid and aldehyde have been known for a long time.

SIMPLE BENZENE COMPOUNDS WITH TWO CARBON CHAINS

We now come to a group of benzenoid metabolic products, three of which seem to me to be of considerable importance, since, as will become clear later, they appear to represent a stage in the biosynthesis of a number of more complex mould metabolites. The three metabolites to which I wish to draw your attention are 6-hydroxy-2-methylbenzoic acid (XXXVIII) from *P. griseo-fulvum*^(60, 58) and *P. flexuosum*⁽⁶¹⁾,

3:5-dihydroxyphthalic acid (XXXIX) from *P. brevi-compactum* (62, 63), and ustic acid (XL) from *Aspergillus ustus* (64). 6-Hydroxy-2-methylbenzoic acid is clearly related to mellein (XLI) from *A. melleus* (65, 66) and from *A. ochraceus* (67, 68), and is, in fact, produced from it by potash fusion. Mellein itself is the lactone of the acid given in

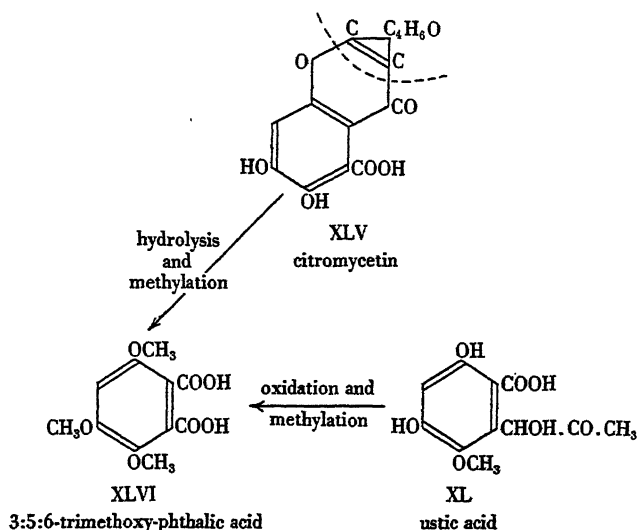


brackets. 3:5-Dihydroxyphthalic acid is also clearly related to the other metabolic products of *Penicillium brevi-compactum*, namely, the acids C₁₀H₁₀O₅ (XLII), C₁₀H₁₀O₆ (XLIII) and C₁₀H₁₀O₇ (XLIV) (62, 69, 70). Each of these three compounds, as well as mellein and ustic acid, has a side-chain of three carbon atoms in varying degrees of oxidation. Ustic acid itself may be regarded as the methoxy derivative of C₁₀H₁₀O₆, although it is not yet established with certainty in which of the possible tautomeric forms the three carbon side-chain is present.

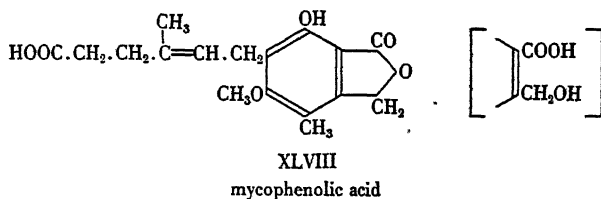
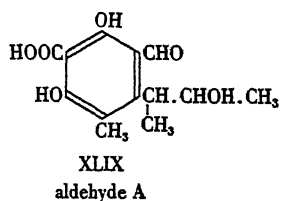
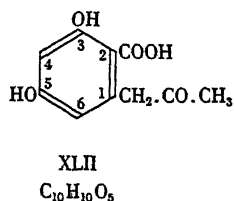
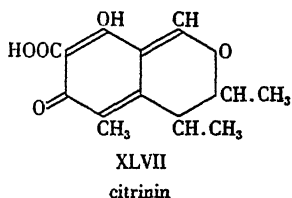
CITROMYCETIN

An interesting example of interrelationship in structure is afforded by the mould metabolite citromycetin (XLV).

Citromycetin is a yellow crystalline substance obtained from cultures of a number of species or strains in the *Penicillium frequentans* group (71), some of which were formerly regarded as species of *Citromyces*—hence the name of the metabolite. The complete structure of citromycetin has not yet been established, but a partial structure, in which the nature of the C₄H₆O group is at present doubtful, is represented by XLV. Alkaline hydrolysis of the dimethyl ether of citromycetin, followed by methylation of the hydrolysis product, yields 3:5:6-trimethoxyphthalic acid (XLVI) (64). The same compound is formed by the oxidation, followed by methylation of the oxidation product, of the metabolite ustic acid (XL) from *Aspergillus ustus* (64).



CITRININ, MYCOPHENOLIC ACID



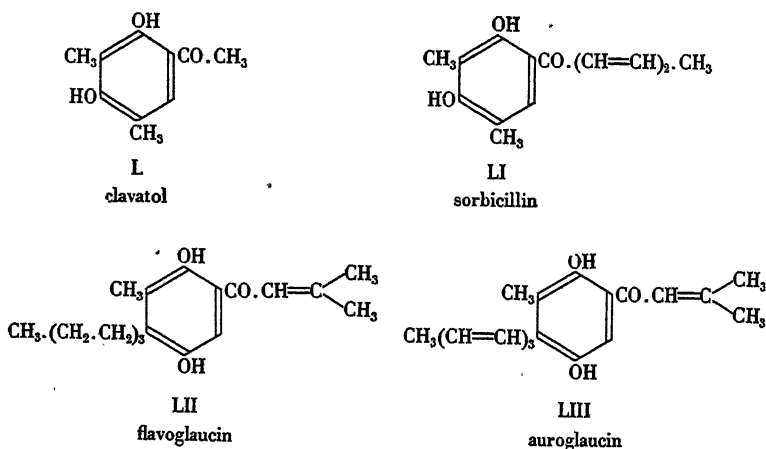
Further examples of interrelationships in structure are afforded by the mould metabolic products citrinin, mycophenolic acid and one of the C_{10} acids, $C_{10}H_{10}O_5$, from *Penicillium brevi-compactum*. Citrinin (XLVII), which is a beautifully crystalline yellow metabolite of *P. citrinum* (72, 73), *Aspergillus terreus* (74) and other moulds, has recently been synthesized by Professor Alexander Robertson and his co-workers at Liverpool (75). The final step in this synthesis was the ring closure of the aldehyde A (XLIX) which, by loss of water and molecular rearrangement, led to the quinonoid body citrinin.

Mycophenolic acid (XLVIII) was first isolated by Alsberg & Black in 1913 from cultures of *Penicillium stoloniferum* (76). Since then it has been reported from a large number of other species and strains in the *P. brevi-compactum* series (62). Mycophenolic acid has not yet been synthesized, but there is strong evidence in support of the structure advanced and shown, in brackets, with the phthalide ring opened (77, 78).

Turning now to the acid $C_{10}H_{10}O_5$ (XLII) from *P. brevi-compactum* it will be seen that by the introduction of a methyl group into position 6 and a carboxyl group into position 4 we arrive at a structure very similar to that of the aldehyde A from which citrinin was derived. Similarly, by the introduction of a methyl group again into position 6 and the appropriate long side-chain again into position 4 we arrive at a structure very closely related to mycophenolic acid.

There is also a general though not complete similarity in structure between the aldehyde A and two other mould metabolites, clavatol and sorbicillin, which have been described quite recently.

CLAVATOL, SORBICILLIN, FLAVOGLAUCIN, AUROGLAUCIN



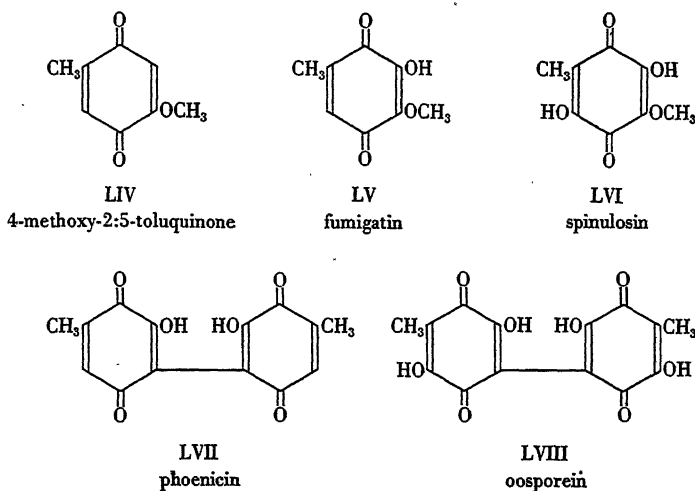
Clavatol (L) was isolated by Bergel *et al.* from culture filtrates of *Aspergillus clavatus* (13). Its structure was determined and its synthesis carried out by Dr C. H. Hassall and Professor A. R. Todd at Cambridge (79).

Sorbicillin (LI) was described by D. J. Cram in America as an orange crystalline metabolic product of *Penicillium notatum*, and was isolated by him from commercial clinical penicillin (80). Hence the possibility is not excluded that sorbicillin is an artefact arising during the manufacturing processes used for the production of penicillin.

Clavatol and sorbicillin may be regarded as the resorcinol analogues of the substituted quinol metabolites flavoglaucin and auroglaucin, both of which have been isolated from the mycelium of a large number of species in the *Aspergillus glaucus* series (81). The general nature of the structures of the yellow flavoglaucin and the orange-red auroglaucin as substituted quinols was established in the Dyson Perrins Laboratory, Oxford (82, 83), although the exact position and nature of the

substituent side-chains was not determined with certainty. Quite recently Italian workers (84) have confirmed the general correctness of this work and have produced evidence that flavoglaucin has structure LII and auroglaucin structure LIII. Flavoglaucin and auroglaucin are thus clearly derivatives of toluquinol, so that we proceed logically to the consideration of mould metabolic products which are derivatives of toluquinone.

DERIVATIVES OF TOLUQUINONE



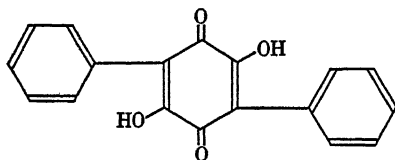
In recent years moulds have proved to be very fruitful sources of a number of different types of quinones. Many moulds are highly coloured—yellow, orange, red, purple and violet—and these colours have frequently been shown to be due to the presence of quinones or their salts, depending on the pH of the medium.

The first group of quinones which I wish to discuss is the series of five derivatives of toluquinone given above. All of them have been synthesized and they form a perfect series of structurally interrelated metabolites.

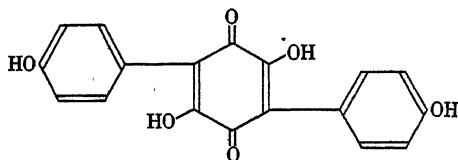
4-Methoxy-2:5-toluquinone (LIV) which is orange yellow in colour was reported a few months ago by workers in the New York Botanical Gardens as a metabolite of laboratory cultures of the higher fungi *Coprinus similis* and *Lentinus degener* (85). Fumigatin (LV), maroon red in colour, occurs in the culture filtrates of an unusual but authentic strain of *Aspergillus fumigatus* (86). It is clearly 3-hydroxy-4-methoxy-2:5-toluquinone (87). Spinulosin (LVI), deep purple black in colour, was first isolated from cultures of *Penicillium spinulosum* (88), and later from a different strain of *Aspergillus fumigatus* (89) from that which gave fumigatin, and quite recently from *Penicillium cinerascens* (26). Spinulosin is clearly 6-hydroxyfumigatin (86, 90).

Phoenicin (LVII) and oosporein (LVIII) are derivatives of 4:4'-ditoluquinone. It will be seen that phoenicin bears the same relationship to fumigatin as does oosporein to spinulosin. Phoenicin was discovered in 1933 by Dr E. Friedheim in Switzerland as a metabolic product of *P. phoeniceum* (91, 92), and its constitution and synthesis were described by Dr T. Posternak (93) who also isolated it from

P. rubrum (94). Oosporein was obtained by Kögl & van Wessem in Holland in 1944 from laboratory cultures of the mould *Oospora colorans* (95).



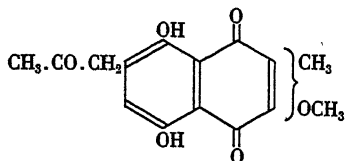
LIX
polyporic acid



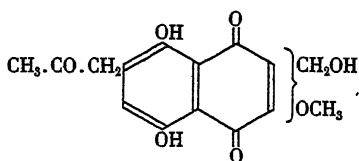
LX
atromentin

Spinulosin (LVI) is also closely related to a number of other naturally occurring derivatives of benzoquinone, and in particular to two substances investigated by Kögl and his co-workers and isolated by them from naturally occurring specimens of two of the higher fungi. These fungal metabolic products are polyporic acid (LIX) from *Polyporus nidulans* (96), and atromentin (LX) from *Paxillus atrotomentosus* (97). It will be seen that the methyl and methoxy groups in spinulosin are replaced by two phenyl groups in polyporic acid, and by two *para*hydroxyphenyl groups in atromentin.

DERIVATIVES OF NAPHTHAQUINONE



LXI
javanicin

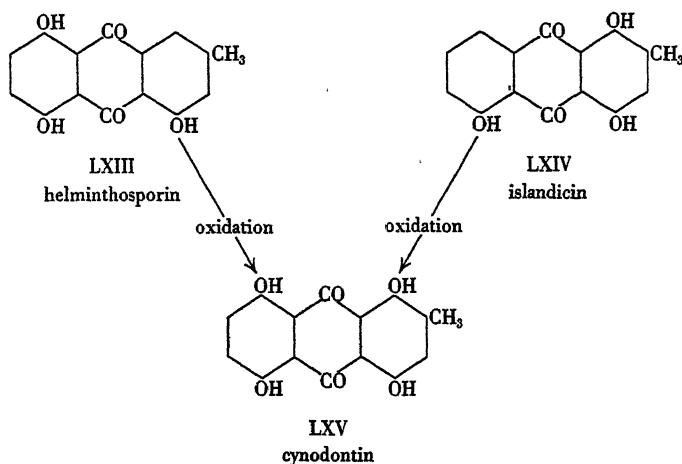


LXII
oxyjavanicin

So far as I have been able to ascertain only two substituted naphthaquinones have been isolated from cultures of moulds, namely, javanicin, $C_{15}H_{14}O_6$ (LXI), and oxyjavanicin, $C_{15}H_{14}O_7$ (LXII). These two substances were prepared from the

blood-red culture filtrate of *Fusarium javanicum* by Dr A. H. Cook and his colleagues at Imperial College, London (98, 99). The molecular structure of these compounds has not yet been completely established with certainty, but, as a result of analytical work and the determination of absorption spectra, javanicin is believed to have the partial structure given above. Oxyjavanicin would then have a similar structure except that a hydroxymethyl group would take the place of the methyl group in javanicin. In the absence of complete molecular structures one can only remark that javanicin contains a methyl and methoxy group as do fumigatin (LV) and spinulosin (LVI), and a $\text{CH}_3\text{CO} \cdot \text{CH}_2$ side-chain as does the $\text{C}_{10}\text{H}_{10}\text{O}_5$ acid (XLII) from *Penicillium brevi-compactum*.

DERIVATIVES OF ANTHRAQUINONE



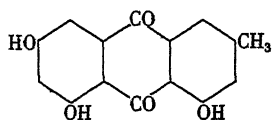
Since 1933 a considerable number of polyhydroxyanthraquinones, previously unknown to organic chemistry, have been shown to be metabolites of different species in a number of different genera of moulds. They are often produced in considerable amounts. Thus the dried mycelium of *Helminthosporium gramineum* contained 30 % of its weight of a mixture of polyhydroxyanthraquinones (100).

The first one which we isolated was helminthosporin (LXIII) from the mycelium of *H. gramineum* (100), *H. catenarium* and *H. tritici-vulgaris* (101) and *H. cynodontis* (102), and I welcome the opportunity of acknowledging publicly my indebtedness to our President for his guidance and collaboration in elucidating the structure of many of the anthraquinones which I shall describe, since he was associated with this work from the earliest days of the isolation of helminthosporin.

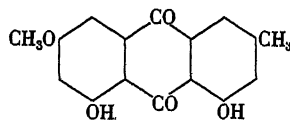
Islandicin (LXIV) was isolated from cultures of *Penicillium islandicum* (103), and cynodontin (LXV) from *Helminthosporium cynodontis* (102), *H. euschlaenae* (102) and *H. avenae* (101). Helminthosporin (104, 105) and cynodontin (106) have been synthesized.

Helminthosporin and islandicin are both readily converted into cynodontin by oxidation with manganese dioxide and concentrated sulphuric acid, so that the close structural relationship between these three substances is evident. Islandicin is also closely related to catenarin (LXXII) and erythroglauicin (LXXIII), which are respectively 7-hydroxy- and 7-methoxy-islandicin.

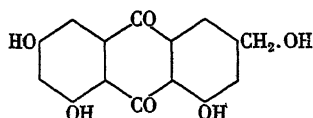
An outstanding feature of many of the anthraquinone mould metabolites is their close structural relationship to *Frangula* emodin. So far as I am aware emodin itself (LXVI) has not been found in any mould, though Kögl isolated it in 1925 from natural specimens of the higher fungus *Dermocybe sanguinea* (107). Physcion (LXVII), the 7-methyl ether of emodin, has long been known as a lichen constituent (108, 109) and has also been isolated from laboratory cultures of seventeen species or strains in the *Aspergillus glaucus* series (82, 110, 111). The previously undescribed ω -hydroxyemodin (LXVIII) was isolated independently and almost simultaneously by ourselves from *Penicillium cyclopium* (112), and by Posternak from *P. citreo-roseum* and was called by him citreorohein (LXVIII) (113, 114). The 4-methyl ether of



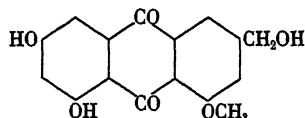
LXVI
emodin



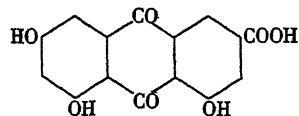
LXVII
physcion



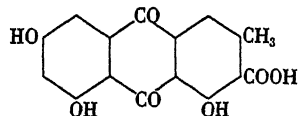
LXVIII
 ω -hydroxyemodin
citreorohein



LXIX
carviolin
roseopurpurin



LXX
emodic acid

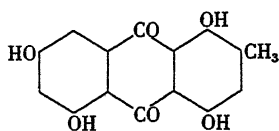


LXXI
endocrocin

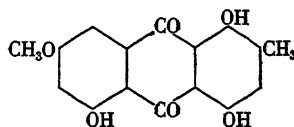
ω -hydroxyemodin was also described independently and almost simultaneously by Hind from *P. carmino-violaceum* (115, 116) and called by him carviolin (LXIX), and by Posternak from *P. roseopurpureum* (117, 118) and called by him roseopurpurin (LXIX). Emodic acid (LXX) occurs along with ω -hydroxyemodin in cultures of *P. cyclopium* (112). Endocrocin (LXXI), which is 3-carboxyemodin, was obtained by Asahina & Fuzikawa from the Japanese lichen, *Nephromopsis endocrocea* (119). It will be noticed that all these six compounds may be regarded as being derived from 3:5-dihydroxyphthalic acid (XXXIX), which, as I have shown previously, is a metabolite of *Penicillium brevi-compactum* (83), and that endocrocin might, on paper at any rate, be quite easily synthesized by coupling 3:5-dihydroxyphthalic acid with the metabolite of *P. griseo-fulvum*, 6-hydroxy-2-methylbenzoic acid (XXXVIII) (80, 81).

Three other anthraquinones, each of which may be regarded as a derivative of emodin, and each of which contains four nuclear hydroxyl groups, are shown on p. 157.

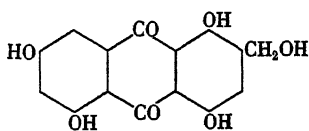
Catenarin (LXXII) from *Helminthosporium catenarium* (101), *H. gramineum* (100), *H. velutinum* and *H. tritici-vulgaris* (101) is clearly 1-hydroxyemodin (120). Erythroglaucin (LXXIII), from fifteen species or strains in the *Aspergillus glaucus* series (111), is clearly the 7-methyl ether of catenarin (120). The molecular constitutions of catenarin and erythroglaucin have been confirmed by synthesis (121, 120), but that of tritisorin (LXXIV) from *Helminthosporium tritici-vulgaris*, although quite probable, is less firmly established (101). Accepting it, however, tritisorin becomes 1-hydroxy- ω -hydroxyemodin.



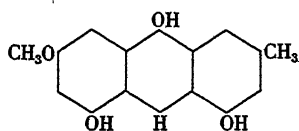
LXXII
catenarin



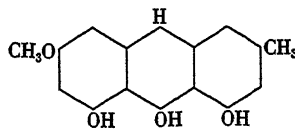
LXXIII
erythroglaucin



LXXIV
tritisorin



LXXV
physcion anthranol A



LXXVI
physcion anthranol B

Catenarin, erythroglaucin and tritisorin can obviously be built up from the mould metabolite 3:5-dihydroxyphthalic acid (XXXIX) (63), coupled, in the case of tritisorin, with a second mould metabolite, gentisyl alcohol (XXXVII) (59), and in the other two cases with toluquinol, which while not yet described as a mould metabolite may be confidently expected to be one in view of its close relationship to gentisyl alcohol.

The close relationship to physcion and hence to emodin of the two anthranols of physcion, A (LXXV) and B (LXXVI), which come from a few species in the *Aspergillus glaucus* series (111), is obvious.*

A number of mould-colouring matters, steadily increasing in number and having high melting-points, are almost certainly polyhydroxy di-anthraquinones, and some of them are clearly derivatives of emodin. For that reason I think that they are

* It will be seen that, without exception, all the anthraquinone metabolites which I have described are derivatives of chrysophanol (chrysophanic acid), 4:5-dihydroxy-2-methylantraquinone. The observation, made by Mr B. H. Howard in my department since this lecture was delivered, that chrysophanol is one of the colouring matters of *Penicillium islandicum* is therefore of considerable interest.

worthy of consideration here, in spite of the fact that their molecular constitution has not yet been completely established.

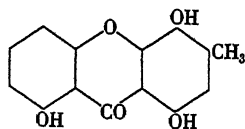
Penicillipsin from the mould *Penicillipsis clavariaeformis* (122) has the empirical formula $C_{30}H_{24}O_8$ and melts at about $330^\circ C$. On thermal decomposition it yields emodin anthranol, and on oxidation with nitric acid it gives tetranitro-emodin.

A pigment from *Penicillium islandicum*, *P. wortmanni* and *P. rugulosum* which we are investigating at present has the probable empirical formula $C_{31}H_{20}O_{11}$, and does not melt below $390^\circ C$. On treatment with cold aqueous sodium hyposulphite, $Na_2S_2O_4$, in sodium carbonate solution it gives 90 % of the theoretical yield of two molecules of emodin, a curious if not indeed an almost unique reaction.

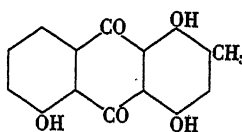
Rugulosin, $C_{22}H_{20}O_8$ (?), m.p. $294^\circ C$, from a large number of different strains of *P. rugulosum*, is not itself an anthraquinone, but is, we believe, a complex and highly substituted benzoyl benzoic acid (unpublished observations). However, it appears to be related to the anthraquinones since, on thermal decomposition, it gives about 30 % of a mixture of emodin and chrysophanol.

Finally, aurofusarin from *Fusarium culmorum* (123), has the empirical formula $C_{30}H_{20}O_{12}$ and does not melt below $360^\circ C$. Little is known of its constitution except that it contains two methoxy groups and that its general properties are those of a polyhydroxy di-anthraquinone.

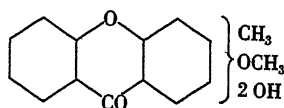
DERIVATIVES OF XANTHONE



LXXVII
ravenelin



LXIV
islandicin



LXXVIII
rubrofusarin

Two xanthone derivatives are known as mould metabolites.

Ravenelin (LXXVII) was isolated from the mycelium of laboratory cultures of *Helminthosporium ravenelii* and *H. turcicum* (124). Its molecular structure is shown above, and this was subsequently confirmed by its synthesis in 1944 by Mull & Nord (125). Its very close structural relationship to the anthraquinone metabolite islandicin (LXIV) from *Penicillium islandicum* is obvious.

Rubrofusarin (LXXVIII) occurs along with the di-anthraquinone derivative aurofusarin in the mycelium of *Fusarium culmorum* (123). It is the monomethyl ether of a methyl trihydroxyxanthone which is isomeric but not identical with ravenelin. Its molecular constitution has not yet been established with certainty, although Mull & Nord suggest two highly speculative structures (125).

The last group of mould metabolites which I wish to discuss is chosen not so much because the members of it are structurally related as because they all contain chlorine.

TABLE 3. MOULD METABOLITES CONTAINING CHLORINE

metabolite	empirical formula	species of micro-organism	investigators	reference
caldariomycin	$C_5H_8O_2Cl_2$	<i>Caldariomyces fumago</i>	London School of Hygiene and Tropical Medicine	(126)
chloromycetin	$C_{11}H_{12}O_5N_2Cl_2$	<i>Streptomyces venezuelae</i>	Parke, Davis Labs.	(127), (128)
erdin	$C_{16}H_{10}O_7Cl_2$	<i>Aspergillus terreus</i>	L.S.H.T.M.	(129)
geodin	$C_{17}H_{12}O_7Cl_2$	<i>Aspergillus terreus</i>	L.S.H.T.M.	(129)
griseofulvin	$C_{17}H_{17}O_6Cl$	<i>Penicillium griseo-fulvum</i> <i>P. janczewskii</i>	L.S.H.T.M. Brian, Curtis & Hemming	(130) (131), (132), (133)
sclerotiorine	$C_{20}H_{19(21)}O_5Cl$	<i>P. sclerotiorum</i>	Curtin & Reilly	(134)
aureomycin	$C_{22}H_{27}O_8N_2Cl(?)^*$	<i>Streptomyces aureofaciens</i>	Lederle Labs.	(135), (136)

* This empirical formula has not been published but is one of the possible formulae calculated from data given in reference (136).

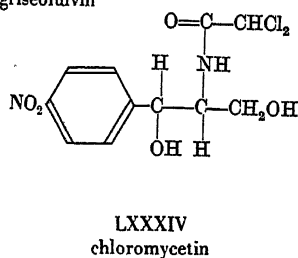
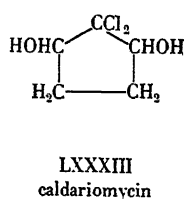
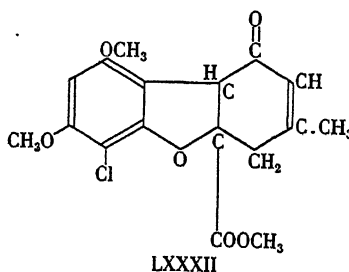
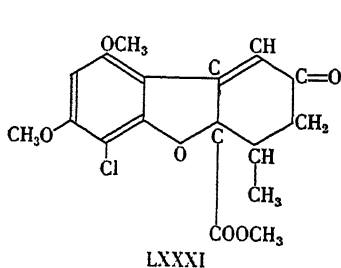
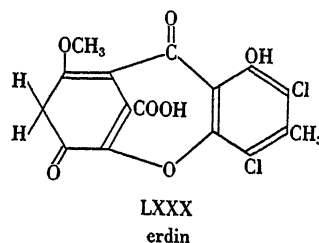
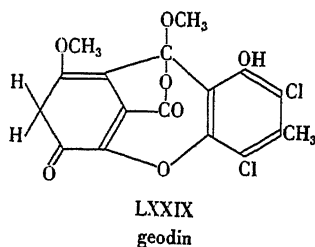
It will be remembered that one of our standard media—Czapek-Dox medium (p. 142)—contains, in addition to glucose, only mineral salts which include half a gram of potassium chloride per litre of medium. A quantitative survey, carried out in the late 1930's, of the chlorine metabolism of 139 species or strains of moulds grown on this medium, revealed the fact that certain species metabolize the inorganic chlorine ion and convert it into non-ionic chlorine (126). This observation, among others which we had made previously, led us to the isolation of four new metabolic products containing organically bound chlorine, namely, caldariomycin from *Caldariomyces fumago*, erdin and geodin from *Aspergillus terreus*, and griseofulvin from *Penicillium griseo-fulvum*. These were followed by the isolation in other laboratories of sclerotiorine from *P. sclerotiorum*, and, quite recently, by chloromycetin and aureomycin from two different species of *Streptomyces*.

All of these seven metabolites are crystalline substances, and chloromycetin and aureomycin, both of which contain nitrogen in addition to chlorine, are the new antibiotics which are unique in having very powerful anti-rickettsial and anti-viral properties. Chloromycetin is the only one which has been synthesized (137), but structural formulae have been advanced for all of them except sclerotiorine and aureomycin. These structural formulae are given below.

Geodin (LXXIX) and erdin (LXXX) are closely related structurally to each other (138, 139), and, thanks to a suggestion of Professor A.R. Todd and Dr Hassall, we now regard geodin, which is optically active, as the pseudo-methyl ester of erdin, which is optically inactive (140).

Two structural formulae (LXXXI and LXXXII) have been proposed for griseofulvin. The original structure (LXXXI) was advanced by ourselves (130), while structure LXXXII was proposed later by Grove & McGowan (133).

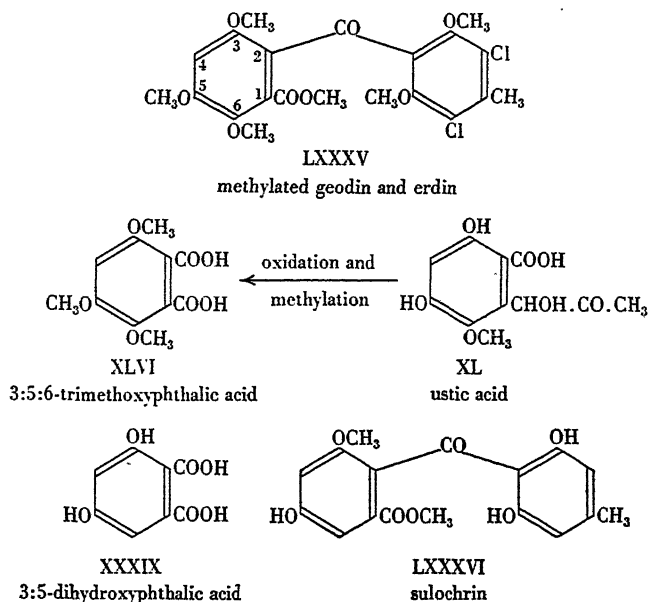
Caldariomycin (LXXXIII) was, until recently, the only chlorine containing metabolite having two chlorine atoms attached to the same carbon atom (126). A similar grouping is now known to be present in chloromycetin (LXXXIV) which, on acid or alkaline hydrolysis, yields, *inter alia*, dichloroacetic acid. Chloromycetin is unique among microbial metabolites and also I believe among any natural products in containing a nitrobenzene grouping. The antibiotic is the D-(-)-threo form (137).



Geodin (LXXIX) and erdin (LXXX) are closely related structurally to at least three other mould metabolites. On methylation with methyl sulphate and alkali, the seven-membered ring opens and there is formed from both geodin and erdin the same highly substituted benzoyl benzoic acid shown in structure LXXXV (140). This compound can clearly be derived from 3:5:6-trimethoxyphthalic acid (XLVI) which is itself formed by the oxidation and methylation of the mould metabolite ustic acid (XL) from *Aspergillus ustus* (84), and by hydrolysis and methylation of citromycetin (XLV) from *Penicillium frequentans* (71).

Further, by removal of the two chlorine atoms in the right-hand ring and of the methoxy group in position 6 in the left-hand ring, we arrive at the same basic

structure as is present in sulochrin (LXXXVI), a metabolic product of the mould *Oospora sulphurea-ochracea* described by Nishikawa (141, 142, 143), and sulochrin in its turn may be derived from 3:5-dihydroxyphthalic acid (XXXIX), a metabolite of *Penicillium brevi-compactum* (63).



BIOLOGICAL SIGNIFICANCE OF MOULD METABOLIC PRODUCTS

The large number of mould metabolic products and, from the chemical point of view, their great variety in structure naturally prompts the question as to what function they serve in the economy of the organisms which produce them.

In trying to answer this question, though in a negative sense, I wish to emphasize that it is my opinion that they cannot, in general, be regarded as by-products of no particular biological importance. If this were so why should many of them be produced in such considerable yields? Further, it should be remembered that these large yields can often only be obtained by harvesting the cultures before all the glucose in the medium has been completely metabolized. Otherwise long-continued incubation leads to their complete disappearance by oxidation to carbon dioxide and water.

There is a growing belief that certain types of metabolites play some part in the oxidation-reduction mechanisms of the moulds which produce them. Thus the culture filtrates of *Aspergillus fumigatus* from which fumigatin (LV) was isolated also contain the corresponding quinol, 3-hydroxy-4-methoxytoluquinol (86). Similarly phenicin (LVII) is accompanied by its leuco derivative, tetrahydrophenicin, in cultures of *Penicillium rubrum* (94); physcion (LXVII) occurs in the mycelium of certain species in the *Aspergillus glaucus* series along with its two reduction products, i.e. the anthranols 4:5-dihydroxy-7-methoxy-2-methyl-9-anthranol (LXXV) and 4:5-dihydroxy-7-methoxy-2-methyl-10-anthranol (LXXXVI) (111); the mycelium of

Helminthosporium leersii contains the yellow luteoleersin, $C_{26}H_{38}O_7$, and its colourless reduction product alboleersin, $C_{26}H_{40}O_7$, of at present undetermined molecular constitution (144); and culture filtrates of the mould *Oidiodendron fuscum* contain, in addition to the orange crystalline fuscin, $C_{15}H_{16}O_5$, its leuco derivative dihydro-fuscin, $C_{15}H_{18}O_5$ (145). All these quinonoid metabolic products and their corresponding leuco derivatives have been shown to be readily inter-convertible *in vitro*, and Friedheim has shown that traces of phenicin increase by 200 to 300% the respiration of unpigmented and washed cells of *Bacillus pyocyaneus* (91, 92).

It seems probable that some of the mould metabolites already described and certainly others which await isolation play a fundamental role in the biological phenomenon known as antibiosis. Thus in addition to those antibiotics of established clinical importance such as penicillin, streptomycin, aureomycin and chloromycetin, a large number of other mould metabolites exhibit very powerful antibacterial and, significantly, antifungal properties *in vitro*. This subject has been adequately reviewed recently by Oxford (146), Birkinshaw (147) and Kavanagh (148). The general biological significance, as distinct from the medical importance, of this aspect of the subject is indicated by the following examples.

(a) Van Luijk (149) showed in 1938 that sterilized culture filtrates of the saprophytic mould *Penicillium expansum* totally inhibited the growth *in vitro* of the plant pathogenic fungus *Pythium de Baryanum* at dilutions of 1:1,280, and he obtained good results in suppressing the attack of the *Pythium* on lucerne seedlings by treating infected soil with the active filtrate. The active principle responsible for this effect, patulin (V), was isolated by ourselves from *Penicillium expansum* in 1943 (150, 151) and independently by Dutch workers from a different strain of the same species (16, 17). They named it expansin. It inhibits completely the growth of a number of species of *Pythium* at a concentration of about 1:500,000 (150).

(b) Rayner (152) and Neilson Jones (153) have shown that the growth failure of conifers at Wareham Heath, Dorset, is associated with the development in the soil of a definite toxicity, of biological origin, to the mycorrhizal fungi normally associated with these trees. Brian and his colleagues (154) have demonstrated that moulds which are commonly abundant in meadow or arable soils, e.g. *Mucor* spp., *Trichoderma viride*, *Penicillium chrysogenum* and *Fusarium* spp., are virtually absent in Wareham Heath soil, in which the mould flora is almost restricted to three groups of *Penicillium* spp., viz. strains of *P. janczewskii*, *P. terlikowskii* and strains in the *P. nigricans-janczewskii* series. They have also shown that each of these three groups gives rise to metabolic products which have been isolated in a state of purity and which, at a very low concentration, have marked biological effects on other fungi. Thus, *P. janczewskii* was shown to produce the chlorine containing metabolite griseofulvin (LXXXI and LXXXII) to which they gave the name 'curling factor' since, at concentrations as low as $1\mu\text{g./ml.}$ it causes excessive branching and distortion of the germ-tubes and hyphae of *Botrytis allii* (131, 132, 133); *P. terlikowskii* produces gliotoxin, $C_{13}H_{14}O_4N_2S_2$, first isolated as a colourless crystalline metabolite of *Trichoderma viride* by Weindling & Emerson (155) which, in addition to being actively antibacterial, is also highly fungistatic to a number of saprophytic and plant pathogenic fungi (156); and strains in the *Penicillium nigricans-*

janczewskii series yield a new unnamed metabolic product crystallizing in glistening red needles and having the probable molecular formula $C_{10}H_8O_4$. This substance is actively fungistatic and prevents the germination of conidia of *Botrytis allii* at concentrations of $0.4\mu\text{g./ml.}$ at pH 3.5 (157).

CONCLUSION

My purpose in this lecture has been to present to you, although unavoidably in a very condensed and incomplete form, the relationships in chemical structure between members of representative classes of fungal metabolic products. For many years after the commencement of our work no such relationship was evident, and the only obvious result of it was the accumulation of a number of new natural products, which while scientifically interesting in themselves, as I think all natural products must be to the chemist and biochemist, seemed to have no obvious place in the biochemistry of living organisms.

I think that this was inevitable in the early days of the new subject of mycological chemistry, since it dealt with a group of organisms which had been sadly neglected by botanists except for a small number of mycologists, and which were in general regarded as a nuisance by bacteriologists, and were almost completely ignored by chemists and biochemists. However, I think it can now rightly be claimed that there are signs of order and of a general underlying design in the biochemistry of the fungi, and my personal view of the mould metabolic products is that they are slowly taking their places in the intricate jigsaw puzzle of natural products.

Thus the slight variations in chemical structure which I have described between some of the benzenoid and quinonoid metabolic products of species of fungi which are closely related morphologically to each other are very reminiscent of similar structural relationships in the colouring matters of flowers which have been shown so clearly by the work of Sir Robert and Lady Robinson and Miss Scott-Moncrieff. This is so much the case that I am convinced that by the collaboration of the geneticist and the mycological chemist which is already beginning to take shape, similar laws will be found to hold with the fungi as have been demonstrated clearly in other living organisms.

Thus J. B. S. Haldane has frequently expressed the view that while biochemical differences between different species of organisms are often striking, the differences between members of a single species—call them strains or mutants, or what you will—may in some ways be an even more fruitful field. I share Haldane's view and firmly believe that by the careful investigation of the metabolic products of closely related species, or of strains or mutants of a single species of fungus, grown under controlled and comparable conditions, much may be learned of the mechanism of intermediate metabolism of these organisms. This line of approach is free from the serious objections to the methods which have been much more favoured such as the so-called fixation method, for example by the addition of sulphite or dimedon to fix acetaldehyde in alcoholic fermentation, and the method of selectively poisoning parts of the catalytic system by the use of specific enzyme poisons like sodium fluoride and iodoacetic acid.

APPENDIX

*List of species cited in text with authorities*Species of *Aspergillus*

A. clavatus Desmazières
A. fumigatus Fresenius
A. glaucus Link
A. itaconicus Kinoshita
A. melleus Yukawa
A. niger van Tieghem

A. ochraceus Wilhelm
A. oryzae (Ahlburg) Cohn
A. terreus Thom
A. ustus (Bainier) Thom & Church
A. wentii Wehmer

Species of *Penicillium*

P. brevi-compactum Dierckx
P. carmino-violaceum Dierckx
P. charlesii G. Smith
P. chrysogenum Thom
P. cinerascens Biourge
P. citreo-roseum Dierckx
P. citrinum Thom
P. claviforme Bainier
P. crateriforme Abbott
P. cyclopium Westling
P. expansum (Link) Thom
P. flexuosum Dale
P. frequentans Westling
P. griseo-fulvum Dierckx
P. islandicum Sopp
P. janczewskii Zaleski
P. minio-luteum Dierckx

P. nigricans Bainier
P. notatum Westling
P. palitans Westling
P. patulum Bainier
P. phoeniceum van Beyma
P. puberulum Bainier
P. roseo-purpureum Dierckx
P. rubrum Grasberger-Stoll
P. rugulosum Thom
P. sclerotiorum van Beyma
P. spiculisporum Lehman
P. spinulosum Thom
P. stipitatum Thom
P. stoloniferum Thom
P. terrestre Jensen
P. terlikowskii Zaleski
P. wortmanni Klöcker

Species of *Helminthosporium*

H. avenae Eidam
H. catenarium Drechsler
H. cynodontis Magnoni
H. euschlaenae Zimmermann
H. gramineum Rabenhorst

H. leersii Atkinson
H. ravenelii Curtis
H. tritici-vulgaris Nisikado
H. turcicum Passerini
H. velutinum Link

Miscellaneous species

Botrytis allii Munn
Caldariomyces fumago Woronichin
Fusarium culmorum (W. G. Smith) Sacc.
Fusarium javanicum Koorders
Oidiodendron fuscum Robak
Oospora colorans van Beyma
Oospora sulphurea-ochracea van Beyma

Penicillioopsis clavariaeformis Solms-Laubach
Pythium de Baryanum Hesse
Streptomyces aureofaciens Duggar
Streptomyces venezuelae Ehrlich, Gottlieb,
 Burkholder, Anderson & Pridham
Trichoderma viride Persoon ex Fries

Higher fungi

Coprinus similis Berkeley & Broome
Cortinarius (Dermocybe) sanguineus (Wulf.)
 Fries
Fomes officinalis (Villars) Neumann
Lentinus degener Kalchbrenner

Lentinus lepideus Fries
Paxillus atrotomentosus (Batsch) Fries
Polyporus nidulans Pers. ex Fries
Trametes suaveolens (Linn.) Fries

Lichens

Parmelia caperata (Linn.) Acheson

Nephromopsis endocrocea Asahina

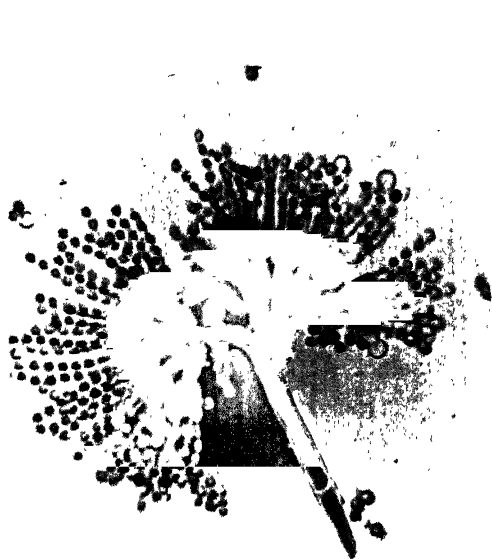


FIGURE 1



FIGURE 2



FIGURE 3



FIGURE 4



FIGURE 5

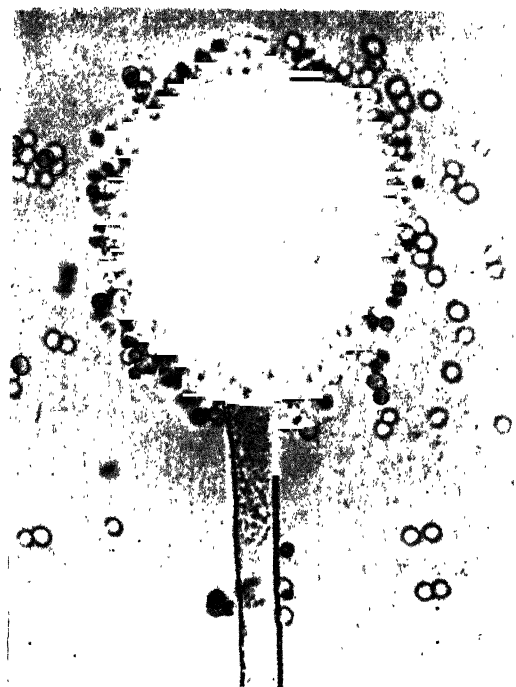


FIGURE 7



FIGURE 6



FIGURE 8



FIGURE 9



FIGURE 10



FIGURE 11

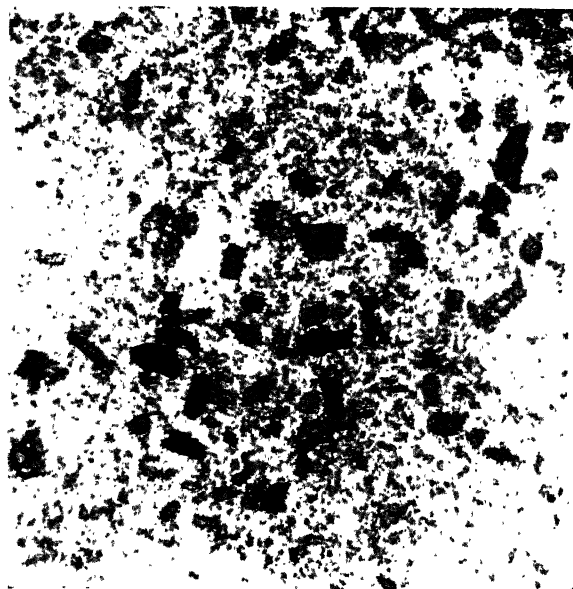


FIGURE 12



FIGURE 13

DESCRIPTION OF PLATES 8 TO 11

I am much indebted to my colleague, Mr George Smith, London School of Hygiene and Tropical Medicine, for the following photomicrographs of some moulds referred to in the text.

PLATE 8

FIGURE 1. *Penicillium brevi-compactum* Dierckx (magn. $\times 600$). Many species in this series produce 3:5-dihydroxyphthalic acid, mycophenolic acid, and the acids $C_{10}H_{10}O_5$, $C_{10}H_{10}O_6$ and $C_{10}H_{10}O_7$.

FIGURE 2. *Penicillium charlesii* G. Smith (magn. $\times 500$). Source of the tetronic acids, γ -methyl-tetronic, carolic, carlic, carolinic and carlosic acids.

FIGURE 3. *Penicillium spinulosum* Thom (magn. $\times 500$). Source of spinulosin.

FIGURE 4. *Penicillium citrinum* Thom (magn. $\times 500$). Source of citrinin.

PLATE 9

FIGURE 5. *Aspergillus chevalieri* (Mangin) var. *intermedius* Thom & Raper (*A. glaucus* series) (magn. $\times 500$). Many species in this series produce the colouring matters flavoglauicin, auroglauicin, erythroglaucin, physcion and the physcion anthranols A and B.

FIGURE 6. *Aspergillus fumigatus* Fresenius (magn. $\times 250$). Source of fumigatin.

FIGURE 7. *Aspergillus flavus* Link (magn. $\times 500$). Source of kojic acid.

FIGURE 8. *Aspergillus terreus* Thom (magn. $\times 250$). Different strains of this species produce terrein, itaconic acid, citrinin, geodin and erdin.

PLATE 10

FIGURE 9. *Fusarium* sp. Typical conidia (magn. $\times 900$).

FIGURE 10. *Helminthosporium monoceras* Drechsler (magn. $\times 90$).

FIGURE 11. *Helminthosporium gramineum* Rabenhorst (magn. $\times 250$). Showing crystals of colouring matters in mycelium. Source of helminthosporin and catenarin.

FIGURE 12. *Penicillium islandicum* Sopp N.R.R.L. 1036 (magn. $\times 250$). Showing crystals of colouring matters in mycelium. Source of islandicin.

PLATE 11

FIGURE 13. *Penicilliopsis clavariaeformis* Solms-Laubach. Showing typical growth in 1 l. flask on a synthetic medium fortified with orange juice. Source of penicilliopsin.

REFERENCES

- (1) Baker, H. 1742 *The microscope made easy*, p. 255. London: R. Dodsley.
- (2) Baker, H. 1743 *An attempt towards a natural history of the polype*, p. 198. London: R. Dodsley.
- (3) Ramsbottom, J. 1941 *Proc. Linn. Soc. Lond.* Session 151, p. 331.
- (4) Wehmer, C. 1891 *Bot. Z.* 49, 233.
- (5) Wehmer, C. 1891 *Ber. dtsh. bot. Ges.* 9, 163.
- (6) Wehmer, C. 1897 *Zbl. Bakt.* II, 3, 102.
- (7) Wehmer, C. 1897 *Chem. Ztg.* 21, 1022.
- (8) Wehmer, C. 1893 *Beiträge zur Kenntnis einheimischer Pilze*. Vol. I. Hannover und Leipzig, Hahn'sche Buchhandlung.
- (9) Wehmer, C. 1893 *Ber. Berlin Akad.* p. 519.
- (10) Raistrick, H. & collaborators 1931 *Studies in the biochemistry of micro-organisms*, Parts 1-18. *Phil. Trans. B*, 220, 1.
- (11) Yabuta, T. 1912 *J. Coll. Agric. Tokyo*, 5, 51.
- (12) Waksman, S. A. 1944 *Science*, 99, 220.
- (13) Bergel, F., Morrison, A. L., Moss, A. R. & Rinderknecht, H. 1944 *J. Chem. Soc.* p. 415.

- (14) Chain, E., Florey, H. W. & Jennings, M. A. 1942 *Brit. J. Exp. Path.* **23**, 202.
- (15) Chain, E., Florey, H. W. & Jennings, M. A. 1944 *Lancet*, **1**, 112.
- (16) Nauta, W. T., Oosterhuis, H. K., Linden, A. C., van Duyn, P. & Dienske, J. W. 1945 *Rec. Trav. chim. Pays-Bas*, **64**, 254.
- (17) Nauta, W. T., Oosterhuis, H. K., Linden, A. C., van Duyn, P. & Dienske, J. W. 1946 *Rec. Trav. chim. Pays-Bas*, **65**, 865.
- (18) Birkinshaw, J. H., Michael, S. E., Bracken, A. & Raistrick, H. 1943 *Lancet*, **2**, 625.
- (19) Sumiki, Y. 1929 *Bull. Agric. Chem. Soc. Japan*, **5**, 10.
- (20) Sumiki, Y. 1931 *J. Agric. Chem. Soc. Japan*, **7**, 819.
- (21) Raistrick, H. & Smith, G. 1935 *Biochem. J.* **29**, 606.
- (22) Clutterbuck, P. W., Raistrick, H. & Reuter, F. 1937 *Biochem. J.* **31**, 987.
- (23) Clutterbuck, P. W., Haworth, W. N., Raistrick, H., Smith, G. & Stacey, M. 1934 *Biochem. J.* **28**, 94.
- (24) Clutterbuck, P. W., Raistrick, H. & Reuter, F. 1935 *Biochem. J.* **29**, 1300.
- (25) Clutterbuck, P. W., Raistrick, H. & Reuter, F. 1935 *Biochem. J.* **29**, 300.
- (26) Bracken, A. & Raistrick, H. 1947 *Biochem. J.* **41**, 569.
- (27) Birkinshaw, J. H. & Raistrick, H. 1936 *Biochem. J.* **30**, 2194.
- (28) Clutterbuck, P. W., Raistrick, H. & Reuter, F. 1935 *Biochem. J.* **29**, 871.
- (29) Alsberg, C. L. & Black, O. F. 1913 *Bull. U.S. Bur. Pl. Ind.* no. 270.
- (30) Birkinshaw, J. H., Oxford, A. E. & Raistrick, H. 1936 *Biochem. J.* **30**, 394.
- (31) Raphael, R. A. 1948 *J. Chem. Soc.* p. 1508.
- (32) Geiger-Huber, M. & Galli, H. 1945 *Helv. Chim. Acta*, **28**, 248.
- (33) Kinoshita, K. 1931 *Bot. Mag.* **45**, 60.
- (34) Kinoshita, K. 1931 *Acta Phytochim. Tokyo*, **5**, 271.
- (35) Calam, C. T., Oxford, A. E. & Raistrick, H. 1939 *Biochem. J.* **33**, 1488.
- (36) Moyer, A. J. & Coghill, R. D. 1945 *Arch. Biochem.* **7**, 167.
- (37) Clutterbuck, P. W., Raistrick, H. & Rintoul, M. L. 1931 *Phil. Trans. B*, **220**, 301.
- (38) Oxford, A. E. & Raistrick, H. 1934 *Biochem. J.* **28**, 1321.
- (39) Birkinshaw, J. H. & Raistrick, H. 1934 *Biochem. J.* **28**, 828.
- (40) Asano, M. & Kameda, Y. 1941 *J. Pharm. Soc. Japan*, **61**, 80.
- (41) Hesse, O. 1898 *J. prakt. Chem.* **58**, 427.
- (42) Asano, M. & Ohta, Z. 1933 *Ber. dtsch. Chem. Ges.* **66**, 1020.
- (43) Asano, M. & Ohta, Z. 1934 *Ber. dtsch. Chem. Ges.* **67**, 1842.
- (44) Thoms, H. & Vogelsang, J. 1907 *Liebigs Ann.* **357**, 145.
- (45) Birkinshaw, J. H. & Raistrick, H. 1936 *Biochem. J.* **30**, 801.
- (46) Birkinshaw, J. H., Chambers, A. R. & Raistrick, H. 1942 *Biochem. J.* **36**, 242.
- (47) Birkinshaw, J. H. & Raistrick, H. 1932 *Biochem. J.* **26**, 441.
- (48) Barger, G. & Dorrer, O. 1934 *Biochem. J.* **28**, 11.
- (49) Dewar, M. J. S. 1945 *Nature*, **155**, 50.
- (50) Dewar, M. J. S. 1945 *Nature*, **155**, 479.
- (51) Dewar, M. J. S. 1945 *Nature*, **155**, 141.
- (52) Cook, J. W. 1945 *Nature*, **155**, 479.
- (53) Erdtman, H. & Gripenberg, J. 1948 *Nature*, **161**, 719.
- (54) Erdtman, H. & Gripenberg, J. 1948 *Acta chem. Scand.* **2**, 625, 639, 644.
- (55) Cook, J. W. & Somerville, A. R. 1949 *Nature*, **163**, 410.
- (56) Birkinshaw, J. H., Bracken, A. & Findlay, W. P. K. 1944 *Biochem. J.* **38**, 131.
- (57) Birkinshaw, J. H. & Findlay, W. P. K. 1940 *Biochem. J.* **34**, 82.
- (58) Raistrick, H. & Simonart, P. 1933 *Biochem. J.* **27**, 628.
- (59) Birkinshaw, J. H., Bracken, A. & Raistrick, H. 1943 *Biochem. J.* **37**, 726.
- (60) Anslow, W. K. & Raistrick, H. 1931 *Biochem. J.* **25**, 39.
- (61) Oxford, A. E., Raistrick, H. & Simonart, P. 1935 *Biochem. J.* **29**, 1102.
- (62) Clutterbuck, P. W., Oxford, A. E., Raistrick, H. & Smith, G. 1932 *Biochem. J.* **26**, 1441.
- (63) Oxford, A. E. & Raistrick, H. 1932 *Biochem. J.* **26**, 1902.
- (64) Stickings, C. E. 1949 *Biochem. J.* **44**, Proc. xxiii.
- (65) Nishikawa, H. 1933 *Bull. Agric. Chem. Soc. Japan*, **9**, 107.
- (66) Nishikawa, H. 1933 *Bull. Agric. Chem. Soc. Japan*, **9**, 148.
- (67) Yabuta, T. & Sumiki, Y. 1933 *J. Agric. Chem. Soc. Japan*, **9**, 1264.
- (68) Yabuta, T. & Sumiki, Y. 1934 *J. Agric. Chem. Soc. Japan*, **10**, 703.

- (69) Oxford, A. E. & Raistrick, H. 1933 *Biochem. J.* 27, 634.
- (70) Oxford, A. E. & Raistrick, H. 1933 *Biochem. J.* 27, 1473.
- (71) Hetherington, A. C. & Raistrick, H. 1931 *Phil. Trans. B*, 220, 209.
- (72) Hetherington, A. C. & Raistrick, H. 1931 *Phil. Trans. B*, 220, 269.
- (73) Coyne, F. P., Raistrick, H. & Robinson, R. 1931 *Phil. Trans. B*, 220, 297.
- (74) Raistrick, H. & Smith, G. 1935 *Biochem. J.* 29, 606.
- (75) Cartwright, N. J., Robertson, A. & Whalley, W. B. 1949 *J. Chem. Soc.* p. 1563.
- (76) Alsberg, C. L. & Black, O. F. 1913 *Bull. U.S. Bur. Pl. Ind.* no. 270.
- (77) Clutterbuck, P. W. & Raistrick, H. 1933 *Biochem. J.* 27, 654.
- (78) Birkinshaw, J. H., Bracken, A., Morgan, E. N. & Raistrick, H. 1948 *Biochem. J.* 43, 216.
- (79) Hassall, C. H. & Todd, A. R. 1947 *J. Chem. Soc.* p. 611.
- (80) Cram, D. J. 1948 *J. Amer. Chem. Soc.* 70, 4240.
- (81) Gould, B. S. & Raistrick, H. 1934 *Biochem. J.* 28, 1640.
- (82) Raistrick, H., Robinson, R. & Todd, A. R. 1937 *J. Chem. Soc.* p. 80.
- (83) Cruickshank, J. H., Raistrick, H. & Robinson, R. 1938 *J. Chem. Soc.* p. 2056.
- (84) Quilico, A. Panizzi, L. & Mugnaini, E. 1949 *Nature*, 164, 26.
- (85) Anchel, M., Hervey, A., Kavanagh, F., Polatnick, J. & Robbins, W. J. 1948 *Proc. Nat. Acad. Sci., Wash.*, 34, 408.
- (86) Anslow, W. K. & Raistrick, H. 1938 *Biochem. J.* 32, 687.
- (87) Baker, W. & Raistrick, H. 1941 *J. Chem. Soc.* p. 670.
- (88) Birkinshaw, J. H. & Raistrick, H. 1931 *Phil. Trans. B*, 220, 245.
- (89) Anslow, W. K. & Raistrick, H. 1938 *Biochem. J.* 32, 2288.
- (90) Anslow, W. K. & Raistrick, H. 1938 *Biochem. J.* 32, 803.
- (91) Friedheim, E. A. H. 1933 *C.R. Soc. Biol. Paris*, 112, 1030.
- (92) Friedheim, E. A. H. 1938 *Helv. chim. Acta*, 21, 1464.
- (93) Posternak, T. 1938 *Helv. chim. Acta*, 21, 1326.
- (94) Posternak, T. 1939 *C.R. Soc. Phys. Hist. Nat. Genève*, 56, 28.
- (95) Kögl, F. & van Wessem, G. C. 1944 *Rec. Trav. chim. Pays-Bas*, 63, 5.
- (96) Kögl, F. 1926 *Liebigs Ann.* 447, 78.
- (97) Kögl, F. & Becker, H. 1928 *Liebigs Ann.* 465, 211, 243.
- (98) Arnstein, H. R. V., Cook, A. H. & Lacey, M. S. 1946 *Brit. J. Exp. Path.* 27, 349.
- (99) Arnstein, H. R. V. & Cook, A. H. 1947 *J. Chem. Soc.* p. 1021.
- (100) Charles, J. H. V., Raistrick, H., Robinson, R., & Todd A. R. 1933 *Biochem. J.* 27, 499.
- (101) Raistrick, H., Robinson, R. & Todd, A. R. 1934 *Biochem. J.* 28, 559.
- (102) Raistrick, H., Robinson, R. & Todd, A. R. 1933 *Biochem. J.* 27, 1170.
- (103) Howard, B. H. & Raistrick, H. 1949 *Biochem. J.* 44, 227.
- (104) Graves, G. D. & Adams, R. 1923 *J. Amer. Chem. Soc.* 45, 2451.
- (105) Raistrick, H., Robinson, R. & Todd, A. R. 1933 *J. Chem. Soc.* p. 488.
- (106) Anslow, W. K. & Raistrick, H. 1940 *Biochem. J.* 34, 1546.
- (107) Kögl, F. & Postowsky, J. J. 1925 *Liebigs Ann.* 444, 1.
- (108) Hesse, O. 1906 *J. prakt. Chem.* 73, 152.
- (109) Hesse, O. 1912 *Liebigs Ann.* 388, 97.
- (110) Raistrick, H. 1937 *Enzymologia*, 4 (Neuberg-Festschrift II), 76.
- (111) Ashley, J. N., Raistrick, H. & Richards, T. 1939 *Biochem. J.* 33, 1291.
- (112) Anslow, W. K., Breen, J. & Raistrick, H. 1940 *Biochem. J.* 34, 159.
- (113) Posternak, T. 1939 *C.R. Soc. Phys. Hist. Nat. Genève*, 56, 29.
- (114) Posternak, T. & Jacob, J. P. 1940 *Helv. chim. Acta*, 23, 237.
- (115) Hind, H. G. 1940 *Biochem. J.* 34, 67.
- (116) Hind, H. G. 1940 *Biochem. J.* 34, 577.
- (117) Posternak, T. 1940 *Helv. chim. Acta*, 23, 1046.
- (118) Raistrick, H. 1941 *Ann. Rep. Chem. Soc.* 38, 262.
- (119) Asahina, Y. & Fuzikawa, F. 1935 *Ber. dtsch. chem. Ges.* 68, 1558.
- (120) Anslow, W. K. & Raistrick, H. 1940 *Biochem. J.* 34, 1124.
- (121) Anslow, W. K. & Raistrick, H. 1941 *Biochem. J.* 35, 1006.
- (122) Oxford, A. E. & Raistrick, H. 1940 *Biochem. J.* 34, 790.
- (123) Ashley, J. N., Hobbs, B. C. & Raistrick, H. 1937 *Biochem. J.* 31, 385.
- (124) Raistrick, H., Robinson, R. & White, D. E. 1936 *Biochem. J.* 30, 1303.
- (125) Mull, R. P. & Nord, F. F. 1944 *Arch. Biochem.* 4, 419.

- (126) Clutterbuck, P. W., Mukhopadhyay, S. L., Oxford, A. E. & Raistrick, H. 1940 *Biochem. J.* **34**, 664.
- (127) Ehrlich, J., Bartz, Q. R., Smith, R. M., Joslyn, D. A. & Burkholder, P. R. 1947 *Science*, **106**, 417.
- (128) Bartz, Q. R. 1948 *J. Biol. Chem.* **172**, 445.
- (129) Raistrick, H. & Smith, G. 1936 *Biochem. J.* **30**, 1315.
- (130) Oxford, A. E., Raistrick, H. & Simonart, P. 1939 *Biochem. J.* **33**, 240.
- (131) Brian, P. W., Curtis, P. J. & Hemming, H. G. 1946 *Trans. Brit. Mycol. Soc.* **29**, 173.
- (132) McCowan, J. C. 1946 *Trans. Brit. Mycol. Soc.* **29**, 188.
- (133) Grove, J. F. & McGowan, J. C. 1947 *Nature*, **160**, 574.
- (134) Curtin, T. P. & Reilly, J. 1940 *Biochem. J.* **34**, 1419.
- (135) Duggar, B. M. 1948 *Ann. N.Y. Acad. Sci.* **51**, Art 2, 175.
- (136) Broschard, R. W., Dornbush, A. C., Gordon, S., Hutchings, B. L., Kohler, A. R., Krupka, G., Kushner, S., Lefemine, D. V. & Pidacks, C. 1949 *Science*, **109**, 199.
- (137) Controulis, J., Rebstock, M. C. & Crooks, H. M. 1949 *Amer. Chem. Soc. Meeting, Div. Med. Chem.* San Francisco, 29 March, Communication v.
- (138) Clutterbuck, P. W., Koerber, W. & Raistrick, H. 1937 *Biochem. J.* **31**, 1089.
- (139) Calam, C. T., Clutterbuck, P. W., Oxford, A. E. & Raistrick, H. 1939 *Biochem. J.* **33**, 579.
- (140) Calam, C. T., Clutterbuck, P. W., Oxford, A. E. & Raistrick, H. 1947 *Biochem. J.* **41**, 458.
- (141) Nishikawa, H. 1936 *Bull. Agric. Chem. Soc. Japan*, **12**, no. 5, 1.
- (142) Nishikawa, H. 1937 *Bull. Agric. Chem. Soc. Japan*, **13**, no. 1, 1.
- (143) Nishikawa, H. 1939 *Acta Phytochim. Japan*, **11**, 167.
- (144) Ashley, J. N. & Raistrick, H. 1938 *Biochem. J.* **32**, 449.
- (145) Michael, S. E. 1948 *Biochem. J.* **43**, 528.
- (146) Oxford, A. E. 1945 *Ann. Rev. Biochem.* **14**, 749.
- (147) Birkinshaw, J. H. 1948 *Trans. Brit. Mycol. Soc.* **30**, 50.
- (148) Kavanagh, F. 1947 *Advances in Enzymology*, **7**, 461.
- (149) van Luijk, A. 1938 *Meded. phytopath. Lab. Scholten*, **14**, 43. Abstract in *Rev. Appl. Mycology*, 1939, **18**, 46.
- (150) Anslow, W. K., Raistrick, H. & Smith, G. 1943 *J. Soc. Chem. Ind., Lond.*, **62**, 236.
- (151) Smith, G. 1947 *Trans. Brit. Mycol. Soc.* **31**, 136.
- (152) Rayner, M. C. 1934, 1936, 1939, 1941 *Forestry*, **8**, 96; **10**, 1; **13**, 19; **15**, 1.
- (153) Neilson Jones, W. 1941 *J. Agric. Sci.* **31**, 379.
- (154) Brian, P. W., Hemming, H. G. & MacGowan, J. C. 1945 *Nature*, **155**, 637.
- (155) Weindling, R. & Emerson, O. H. 1936 *Phytopathology*, **26**, 1068.
- (156) Brian, P. W. 1946 *Trans. Brit. Mycol. Soc.* **29**, 211.
- (157) Curtis, P. J. & Grove, J. F. 1947 *Nature*, **160**, 574.

Vibration spectra of hydrocarbon molecules

II. Skeletal frequencies in certain branched paraffins

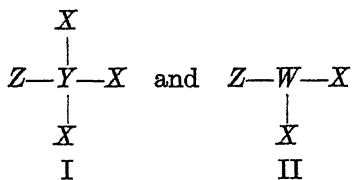
BY DELIA M. SIMPSON, *Newnham College*, AND G. B. B. M. SUTHERLAND, F.R.S.*
Pembroke College, Cambridge

(Received 12 May 1949)

Saturated paraffins containing the terminal grouping $(\text{CH}_3)_3\text{C}...$ show strong absorption bands in the infra-red at 1250 and 1200 cm^{-1} , while those containing the terminal grouping $(\text{CH}_3)_2\text{CH}...$ show strong bands at 1170 and 1145 cm^{-1} . It has been shown that these and other characteristic frequencies should be assigned as vibrations of the carbon skeleton of these structural units. Further, by using a simple valency force field on the model molecules X_3YZ and X_2WZ (where X is a CH_3 group, Y a carbon atom, W a CH group, and Z a large mass representing the remainder of the molecule), satisfactory agreement between observed and calculated frequencies is obtained. The implications of these results in the more general problem of the interpretation of the vibration spectra of hydrocarbons is considered. In particular, it is emphasized that such skeletal frequencies are at least partly responsible for the intense and complex absorption spectra of branched paraffins, in the interpretation of which stress has hitherto been placed mainly on the role of vibrations of the hydrogen atoms.

INTRODUCTION

In an earlier paper (Sheppard & Sutherland 1949) a general classification of the frequencies of hydrocarbon molecules was given and one particular class of frequencies occurring in unsaturated hydrocarbons was discussed in detail. In the present paper attention will be concentrated on saturated paraffins, more particularly those containing terminal groups of the type $(\text{CH}_3)_3\text{C}...$ and $(\text{CH}_3)_2\text{CH}...$ Examination of the infra-red spectra of the paraffins up to and including the nonanes revealed the fact that whenever a paraffin contained a 2, 2-dimethyl grouping it showed two strong absorption bands very close to 1250 and 1200 cm^{-1} , while paraffins containing the 2-methyl grouping showed two strong bands near 1170 and 1145 cm^{-1} . The initial object of the work described in this paper was to give a theoretical interpretation of these two correlations, the general method of approach being to isolate these characteristic groups and to treat their vibrations as virtually independent of the rest of the molecule. To simplify matters still further, the hydrogen vibrations were ignored so that the problem was reduced to considering the vibrations of the idealized molecules I and II,



where Z is a large mass (the rest of the paraffin) X is a CH_3 group, Y a carbon atom, W a CH group and all the angles between the C-C bonds are taken as tetrahedral.

* Now at the Randall Laboratory of Physics, University of Michigan, Ann Arbor, Michigan, U.S.A.

The calculations are based essentially on the assumption of a simple valency force field. It should be emphasized that such a method of treatment can provide only approximate agreement between the absolute values of observed and calculated frequencies, but if it can give a satisfactory justification for the observed correlation of certain frequencies with certain skeletal groups in hydrocarbons, the matter of detailed agreement in the absolute values is really a separate problem which may be regarded as the next step in the interpretation of these highly complex spectra. A preliminary note giving some of the main results of this approach has already appeared (Simpson & Sutherland 1947).

EXPERIMENTAL DATA

Most of the experimental data on which the above correlations are based were accumulated in the course of a programme of research carried out during the war on the infra-red spectra of hydrocarbons in aviation fuel (Sutherland & Thompson 1943; Fellgett, Harris, Simpson, Sutherland, Thompson, Whiffen & Willis 1946). It is not proposed to quote these results in detail as it is planned to publish them separately together with the various correlations that may be derived from them. Valuable confirmatory evidence for these correlations has also been provided by the infra-red data from various laboratories in U.S.A. (*American Petroleum Institute Research Project 44, Reports*, 1943-8) and by the Raman spectra published by Fenske, Braun, Wiegand, Quiggle, McCormick & Rank (1947). Some of the experimental data are quoted later in the paper to illustrate the agreement between calculated and observed frequencies (tables 2, 4 and 7).

SKELETAL MODES OF 2, 2-DIMETHYL PARAFFINS

The model used for the 2, 2-dimethyl paraffins is shown in figure 1 where each X represents a point-mass of 15 (CH_3 group), Y represents a point-mass of 12 (carbon

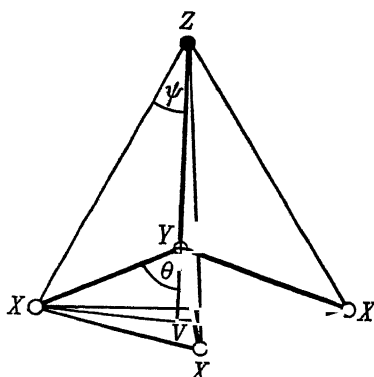


FIGURE 1

atom) and Z represents a point-mass of $N = 14n + 1$ (the alkyl group $\text{C}_n\text{H}_{2n+1}$) situated at the point where the perpendicular from the centre of gravity of the remainder of the hydrocarbon chain intersects the line VY . This last is a rough but

adequate way of representing the effect of variations in the chain length of the alkyl group on the vibrations of the $(\text{CH}_3)_3\text{C}$ group. It will become apparent that all but one of the frequencies are very insensitive to the position of Z along the line YZ . The three bonds YX are assumed to have the usual C-C single bond distance of $1.54 \text{ \AA}$ and to be at the tetrahedral angle ($109^\circ 28'$) with one another and with YZ . ψ is used to denote the angle XZY and θ to denote the angle between XY and YV the perpendicular from Y on to the plane XXX .

The equations for the vibration frequencies of the X_3YZ molecule based on a simple valency force field require two force constants, viz. k for the stretching of YX or YZ , and k' for the deformation of any of the tetrahedral angles between the C-C bonds. They may best be obtained from those given by Slawsky & Dennison (1939) for a more complex field by suitable simplification. Preliminary calculations using average values for k and k' , taken from the literature on some simple hydrocarbon molecules, showed that two frequencies were obtained in the neighbourhood of 1250 and 1200 cm.^{-1} which showed remarkably little variation as Z was varied over a wide range. The values of the force constants were then adjusted to obtain a closer fit with the experimental data. The particular values chosen were $k = 4.65 \times 10^5$ dynes/cm. and $k' = 0.45 \times 10^5$ dynes/cm., and using these, all six skeletal frequencies were calculated for a series of 2, 2-dimethyl paraffins. The results are given in table 1, with the conventional numerical notation generally used for the X_3YZ molecule, viz. ν_1, ν_3, ν_5 are the parallel and ν_2, ν_4, ν_6 the doubly degenerate perpendicular frequencies. Neopentane (2, 2-dimethylpropane) has been included as the first member of the series, although in this case ν_2 and ν_3 coalesce into one triply-degenerate frequency, as do ν_5 and ν_6 .

TABLE 1. CALCULATED SKELETAL FREQUENCIES OF $(\text{CH}_3)_3\text{CZ}$ MOLECULES

Z	N (mass of Z)	parallel frequencies			perpendicular frequencies		
		ν_1	ν_3	ν_5	ν_2	ν_4	ν_6
CH_3	15	727	1288	443	1288	315	443
C_2H_5	29	681	1229	389	1275	295	427
C_3H_7	43	663	1211	362	1262	283	424
C_4H_9	57	657	1203	339	1257	273	422
C_5H_{11}	71	654	1197	324	1254	271	420
C_6H_{13}	85	650	1193	315	1251	255	420
C_7H_{15}	99	648	1191	308	1249	252	420
	∞	646	1177	226	1240	234	420

Examination of table 1 shows that apart from the jump from neopentane to 2, 2-dimethylbutane, ν_2 and ν_3 show no large shifts along this series of paraffins but decrease slowly by only 39 and 97 cm.^{-1} respectively in going from neopentane to 2, 2-dimethylnonane. The corresponding experimental values are given in table 2, from which it appears that ν_2 is sensibly constant in going from neopentane to 2, 2-dimethylheptane while ν_3 changes by nearly 50 cm.^{-1} over the same series. Although both frequencies are fairly constant it is satisfactory that confirmation is also given to the predicted slight variation in ν_3 compared to ν_2 . It should be

TABLE 2. OBSERVED FREQUENCIES OF 2, 2-DIMETHYL PARAFFINS

molecule	frequency numbering (conventional notation)										references
	n_1 ν_1	n_2 ν_4	n_3	n_3	n_3	n_4	n_4	n_4	n_4	n_4	
2, 2-dimethylpropane (neopentane)	R 731 (p)	331 (dp)	1252 (dp)	1252 (dp)	1250	415 (dp)	411	—	—	—	a, b
2, 2-dimethylbutane	IR 714 (0.08)	260 (—)	1221 (0.86)	1258 (0.76)	1250	361 (0.5)	418	—	—	—	c
2, 2-dimethylpentane	IR 710	—	1210 (0.56)	1254 (0.64)	1252	340 (0.56)	—	—	—	—	d
2, 2-dimethylhexane	R 746 (0.08)	740	1204 (0.56)	1251 (0.79)	1250	303 (0.4)	—	—	—	—	e
2, 2-dimethylheptane	IR 730	188	1202 (0.7)	1251 (0.5)	1250	283 (—)	414	—	—	—	f
	R 750 (—)	—	1203	1250	—	—	—	—	—	—	g

Figures in brackets are polarization measurements; — indicates no observations available.

- References: a Rank & Bordner (1935).
 b Sheppard (1948).
 c Sutherland & Thompson (1943).
 d Fenske *et al.* (1947).
 e Avery & Ellis (1942).
 f American Petroleum Institute, Project 44, *Catalogue of infra-red absorption spectral data*, nos. 67, 670.
 g Fellgett *et al.* (1946).

emphasized again at this point that while the approximate method of treatment used here cannot be expected to achieve exact agreement with experimental data, it is just such general trends in the observed spectra which should be predictable.

It also follows from table 1 that the 2, 2-dimethyl paraffins should possess fairly constant frequencies in the neighbourhood of 660 cm.^{-1} (ν_1) and 420 cm.^{-1} (ν_6). The experimental data in table 2 indicate a very constant frequency close to 415 cm.^{-1} which may therefore be assigned to ν_6 but there is no characteristic frequency near 660 cm.^{-1} . On the other hand, there is definite evidence from the Raman spectra of a characteristic frequency near 730 cm.^{-1} which may reasonably be assigned to ν_1 . This assignment has already been made by Sheppard (1948); the numerical disagreement with the calculated value may be attributed to deficiencies in the force field used. Table 2 also includes tentative assignments of observed frequencies to ν_4 and ν_5 , but these must be regarded as much less certain in view of the fact that these frequencies may be expected to vary considerably along the series and in the case of ν_4 to be very sensitive to the particular model chosen, i.e. to the value of ψ . Examination of the available data for some sixteen more complex paraffins having a 2, 2-dimethyl grouping, indicates that their spectra show frequencies which may also be fitted into the same general scheme. This provides further evidence in favour of the assignments in the simple 2, 2-dimethyl paraffins.

In spite of the foregoing general agreement between calculated and observed frequencies for 2, 2-dimethyl paraffins, the fact that alternative assignments have been proposed for some of the characteristic frequencies (Sheppard 1948; Rasmussen 1948) makes it necessary to examine in greater detail the reasons for these assignments and also the basic assumptions underlying the calculations. This is best done with reference to the first member of the series, viz. 2, 2-dimethylpropane or neopentane, since the high degree of symmetry of this particular molecule and the fact that its spectrum has been investigated much more extensively, make it especially suitable for detailed consideration.

NEOPENTANE

The relationships between the four fundamental frequencies of the X_4Y molecule and the six fundamental frequencies of the X_3YZ molecule are well known and are summarized in table 3, together with the selection rules. It is essential that any assignment of the skeletal modes of the 2, 2-dimethyl paraffins (X_3YZ) should be correlated with those of neopentane (X_4Y) in such a way that these fundamental relationships are satisfied. Thus the two characteristic frequencies of the 2, 2-dimethyl paraffins near 1200 and 1250 cm.^{-1} have been assigned in this paper to ν_3 and ν_2 respectively, since they then arise naturally from the splitting up of the triply-degenerate frequency ν_3 of neopentane at 1250 cm.^{-1} . The observed fact that 1250 cm.^{-1} frequency is more constant than the 1200 cm.^{-1} frequency in the 2, 2-dimethyl paraffins is to be expected since the former (ν_2) is a perpendicular frequency of the $(\text{CH}_3)_3\text{C}$ group and so will be much less affected than the latter parallel one (ν_3) when the rest of the paraffin (i.e. the group $\text{C}_n\text{H}_{2n+1}$) varies.

The interpretation proposed above for the regularities in the spectra of the 2, 2-dimethyl paraffins can therefore be checked by a more detailed study of the assign-

ments and of the corresponding calculations for neopentane. The experimental data for this molecule are given in the first two columns of table 4. The assignment of the skeletal modes originally given by Wall & Eddy (1938) and followed by Silver (1940), Pitzer & Kilpatrick (1946), Sheppard (1948) and in the present discussion is given in the third column, while an alternative assignment favoured by Kohlrausch & Köppl (1934), Rank & Bordner (1935), Young, Koehler & McKinney (1947) and by

TABLE 3. RELATIONSHIPS BETWEEN THE FUNDAMENTAL MODES
OF X_4Y AND X_3YZ MOLECULES

T_d sym- metry class	X_4Y (neopentane)				C_{3v} sym- metry class	X_3YZ (2,2-dimethyl paraffins)					
	selection rules					selection rules					
	num- bering	degen- eracy	selection rules			num- bering	degen- eracy	selection rules			
			Raman	infra-red				Raman	infra-red		
A_1	n_1	1	p	ia	A_1	ν_1	1	p	$\downarrow$		
E	n_2	2	dp	ia	E	ν_4	2	dp	$\downarrow$		
T_2	n_3	3	dp	a	A_1	ν_3	1	p	$\downarrow$		
					E	ν_2	2	dp	$\downarrow$		
T_2	n_4	3	dp	a	A_1	ν_5	1	p	$\downarrow$		
					E	ν_6	2	dp	$\downarrow$		

TABLE 4. VIBRATION SPECTRUM OF NEOPENTANE

observed frequencies		assignments		calculated frequencies		
infra-red (vapour) Sutherland & Thompson (1943)	Raman (liquid) Rank & Bordner (1935) Sheppard (1948)			Wall & Eddy (1938)	Ras- mussen (1948)	present
1467	1455 (—)	δ (CH ₃ , CH ₂)	δ (CH ₃ , CH ₂)	—	—	—
1370	—	δ (CH ₃)	δ (CH ₃)	—	—	—
1250	1252 (<i>dp</i>)	<i>n</i> ₃	δ (CH ₃)	1264	—	1258
919	925 (<i>dp</i>)	δ (CH ₃)	<i>n</i> ₃	—	983	—
—	731 (<i>p</i>)	<i>n</i> ₁	<i>n</i> ₁	(732)	729	724
				assumed		
no obser- vations	415 (<i>dp</i>)	<i>n</i> ₄	<i>n</i> ₄	383	422	398
	331 (<i>dp</i>)	<i>n</i> ₂	<i>n</i> ₂	(331)	230	346
				assumed		
force constants used in calculating		<i>k</i>		4.71	4.10	4.60
frequencies (10 ⁵ dynes/cm.)		<i>k'</i>		0.32	0.35	0.35

Rasmussen (1948), is given in the fourth column. The only difference between these two assignments concerns n_3 , the question being whether it is to be associated with the 1250 cm^{-1} or with the 920 cm^{-1} frequency. Since n_3 is the frequency from which the pair of characteristic frequencies of the 2,2-dimethyl paraffins is derived, its assignment is critical.

One of the reasons given by Rasmussen (1948) for assuming $n_3 = 920 \text{ cm}^{-1}$ is because this implies that the 1250 cm^{-1} frequency must then be a rocking mode of the CH_3 group. This argument is really a negative one and is based on the empirical observation that many paraffins with several methyl groups possess frequencies

between 1250 and 1125 cm^{-1} , and on the assumption (made by many workers following Kohlrausch) that these frequencies are too high to arise from carbon skeletal vibrations. However it has been shown in the preceding section that 1250 cm^{-1} is a possible value for a skeletal frequency with quite reasonable values of the force constants. Moreover, the splitting of the frequency at 1250 cm^{-1} in neopentane into two frequencies near 1250 and 1200 cm^{-1} in 2, 2-dimethyl paraffins is accounted for very naturally if 1250 cm^{-1} is assumed a skeletal mode. On the other hand if $n_3 = 920 \text{ cm}^{-1}$, there should be a splitting of this frequency in passing from neopentane to the 2, 2-dimethyl paraffins. The experimental data show that this does not occur consistently in both the Raman and infra-red spectra. Similarly, on Rasmussen's assignment, the 1250 cm^{-1} frequency in neopentane would be expected to be multiple and this is also not observed.

The other principal reason quoted for taking $n_3 = 920 \text{ cm}^{-1}$ is that the assignment is confirmed by calculations based both on a central force field and also on a valency force field. With regard to the former (Kohlrausch & Köppl 1934) it should be noted that the value for the C-C force constant in their central force field calculation is 2.03×10^5 dynes/cm. or less than half the value usually attributed to this force constant. With regard to the latter (Rasmussen 1948), the agreement between calculated and observed frequencies as shown in table 4 is very poor for n_2 . Moreover the value assumed for the C-C force constant is the unusually low one of 4.10×10^5 dynes/cm. and it is clear that this would have to be lowered still further to improve the agreement for his assignment of n_3 . The assignment of $n_3 = 1250 \text{ cm}^{-1}$ is therefore preferred and further justification is sought by more detailed examination of the calculations on the neopentane frequencies.

In the first place, it might be argued that the valency force field used in the preceding section is too crude an approximation and that too much weight should not be put on any agreements between such calculated frequencies and those actually observed. However, it is possible to test the adequacy of the valency force field in a highly symmetrical molecule such as neopentane. Rosenthal (1934*b*) has shown that since only two force constants are used to correlate four frequencies, two relations must exist between the four frequencies and the atomic masses, viz.

$$\left(\frac{n_1 n_2}{n_3 n_4}\right)^2 = \frac{3M}{4m + M}, \quad (1)$$

$$\text{and} \quad n_3^2 + n_4^2 = \frac{n_1^2}{3} \left(2 + \frac{4m + M}{M}\right) + \frac{2}{3} n_2^2 \left(1 + 2 \frac{(4m + M)}{M}\right), \quad (2)$$

where M is the mass of Y and m that of X . Using the assignment in column 3 of table 3, the first equation is satisfied to within 13 % and the second to within less than 0.4 %. Whereas using the assignment of column 4, the corresponding discrepancies are 59 and 41 %. This may be regarded as a further argument in favour of (a) the present assignment, (b) the use of a simple valency force field in connexion with it. Column 7 of table 4 gives the calculated frequencies obtained with a simple valency force field using $k = 4.60 \times 10^5$ dynes/cm. and $k' = 0.35 \times 10^5$ dynes/cm. These calculations were made using the equations of Lechner (1933) which are

more convenient to manipulate for the symmetrical molecule than those of Slawsky & Dennison (1939). The agreement with the observed values is very satisfactory.

It is important nevertheless to get some idea of the size of the interaction terms neglected in the simple valency force field and so it was decided to investigate more general potential functions. The most general potential function of the second degree for an X_4Y tetrahedral pentatomic molecule requires five force constants. Rosenthal (1934*a*) has investigated one form of this function, using general force constants denoted by A , B , C , D and E which have no physical interpretation, while Straley, Tindal & Nielsen (1942) have discussed the general valency force field with force constants K (X - Y stretching), K' (X - Y - X angle deformation) and K_α , K_β and K_γ interaction constants. The relationships between these two equivalent sets of force constants may readily be deduced, and either set computed from the vibrational frequencies. Since the physical interpretation of the second set is more straightforward it will be used in this discussion.

Because neopentane has only four skeletal frequencies, and since no data are available for any symmetrically substituted isotopic molecule, it is necessary to assume a value for one of the force constants. Accordingly values of K were assumed close to that used for the stretching force constant k in the simple valency force field. The results are shown in table 5. In agreement with the indications from the equations (1) and (2), the interaction terms K_α , K_β and K_γ of the more elaborate force field are very small. This treatment shows that although the simple valency force field does not give correct absolute values for the force constants, the values so derived are a very fair approximation.

TABLE 5. FORCE CONSTANTS OF X_4Y MOLECULES (10^5 DYNES/CM.)

neopentane (tetramethylmethane) ($C(CH_3)_4$)					methane (CH_4) tetradeuteromethane (CD_4)			
valency force field		general valency force field			valency force field			
k	4.60	(4.50) assumed	(4.60) assumed	(4.70) assumed	K	4.916	4.878	k
k'	0.35	0.396	0.384	0.380	K'	0.422	0.443	k'
		-0.002	0.051	0.119	K_α	0.289		
		0.141	0.065	0.008	K_β	0.068		
		0.075	0.062	0.059	K_γ	-0.033		

It is interesting to compare the force constants in $C(CH_3)_4$ with the corresponding constants when the general valency force field is applied to CH_4 and CD_4 . Here sufficient data are available to allow them to be determined unambiguously. These were computed from the values of A , B , C , D and E given by Voge & Rosenthal (1936) and are also shown in table 5, together with the values of k and k' for the simple valency force field (Slawsky & Dennison 1939). The parallelism between the five potential constants of methane and tetramethylmethane (neopentane) is very striking.

SKELETAL MODES OF 2-METHYL PARAFFINS

The model used for the 2-methyl paraffins is shown in figure 2 and is analogous to that used earlier for the 2, 2-dimethyl paraffins. X again represents a point-mass of 15 (CH_3 group), W is a point-mass of 13 (CH group), and Z is a point-mass of $N = 14n + 1$ (the alkyl group $\text{C}_n\text{H}_{2n+1}$). Bond lengths WX and WZ are taken to be equal and tetrahedral angles are assumed between C-C bonds. The angle between WV , the perpendicular from W on to the plane XXZ , and WX or WZ is denoted

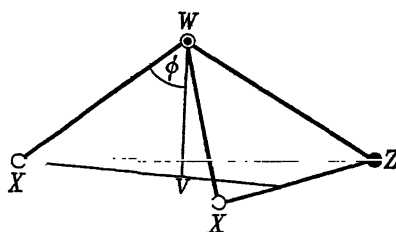


FIGURE 2

by ϕ . Such a molecule has six non-degenerate fundamental skeletal frequencies, but for the first member of the series (*iso*-butane) Z is a methyl group and two pairs of these coalesce to give four fundamentals, two of which are doubly degenerate. The relationships between these two sets of fundamentals are given in table 6.

TABLE 6. RELATIONSHIPS BETWEEN THE FUNDAMENTAL MODES
OF X_3W AND X_2WZ MOLECULES

C_{3v} symmetry class	X_3W (<i>iso</i> -butane)			
	numbering	degeneracy	selection rules	
			Raman	infra-red
A	n_1	1	p	1
A	n_2	1	p	1
E	n_3	2	dp	1
E	n_4	2	dp	1
C_s symmetry class	X_2WZ (2-methyl paraffins)			
	numbering	degeneracy	selection rules	
			Raman	infra-red
A'	ν_1	1	p	a
A'	ν_2	1	p	a
A''	ν_3	1	dp	a
A'	ν_5	1	p	a
A'	ν_4	1	p	a
A''	ν_6	1	dp	a

Again a simple valency force field is assumed and the appropriate equations connecting the fundamentals with the two force constants are most conveniently derived from those of Rosenthal's (1935) more general treatment. The values taken for the

TABLE 7. OBSERVED FREQUENCIES OF 2-METHYL PARAFFINS

molecule	frequency numbering (conventional notation)									
	n_1 ν_1	n_2 ν_2	n_3			n_4				references
X_3W X_3WZ										
2-methylpropane (<i>iso</i> -butane)	R 795 (<i>p</i>)	370 (<i>dip</i>)	1170 (<i>dip</i>)			435 (<i>p</i>)				α
	IR 793	—	—	1170		?		437 422	$b, c(i)$	
2-methylbutane	R 794 (0.2)	368 (—)	1154 (0.7)	1175 (0.6)		?		459 (—)	d	
2-methylpentane	R 815 (0.2)	325 (0.3)	1154 (0.6)	1150 (0.8)	1170	382 (—)	420	442 (—)	$b, c(ii)$	
	IR 817	—	—	1143	1170	?	—	—	b	
2-methylhexane	R 822 (—)	311 (0.2)	1148 (0.6)	1179 (0.7)	1170	?	—	430 (—)	d	
2-methylheptane	IR 814 (—)	289 (—)	1149 (0.4)	1145 (0.4)	1170	317 (—)	—	434 (—)	b	
	IR 840	—	—	1147	1172	?	—	?	d	
2-methyloctane	R 826 (—)	256 (—)	1148 (0.9)	1175 (—)	1170	?	—	—	b	
	IR 820	—	—	1140	1170	—	—	—	e	
2-methylnonane	R —	—	—	—	1167	—	—	—	$c(iii)$	
	IR 833 810	—	—	1142	1167	—	—	?	$c(iii)$	

Figures in brackets are polarization measurements; — indicates no observations available.

References: *a* Ananthakrishnan (1936).

b Sutherland & Thompson (1943).

c American Petroleum Institute, Project 44, Catalogue of infra-red absorption spectral data (i) no. 62; (ii) nos. 64, 555; (iii) nos. 608, 614.

d Fenske *et al.* (1947).

e Fellgett *et al.* (1946).

force constants were 4.50×10^5 dynes/cm. for the C-C stretching (k) and 0.40×10^5 dynes/cm. for the angle deformation (k'), as these gave the closest fit with observed data (table 7). The most characteristic feature in the spectra of the 2-methyl paraffins is the occurrence of two frequencies (strong in both Raman and infra-red spectra) near 1170 and 1145 cm^{-1} , the former being virtually constant in going from *iso*-butane to 2-methylnonane, while the latter varies between 1170 and 1140 cm^{-1} over the same range. It will be observed that the calculations (table 8) lead to a similar pair of frequencies, viz., ν_3 and ν_5 , of which ν_3 remains nearly constant in going from *iso*-butane to 2-methylnonane, while ν_5 decreases through the same series from 1182 to 1093 cm^{-1} . These two frequencies coalesce into one doubly-degenerate vibration ν_3 in *iso*-butane in the same way as the pair of frequencies at 1250 and 1200 cm^{-1} in the 2,2-dimethyl paraffins became one triply-degenerate mode in neopentane. The agreement between the calculated and observed values of ν_3 and ν_5 can be considered adequate in view of the simplifications introduced. As with the 2,2-dimethyl paraffins, the aim at this stage is to account for general features of the spectra.

TABLE 8. CALCULATED SKELETAL FREQUENCIES OF $(\text{CH}_3)_2\text{CHZ}$ MOLECULES

Z	N (mass of Z)	ν_1	ν_2	ν_3	ν_5	ν_4	ν_6
CH_3	15	870	371	1182	1182	396	396
C_2H_5	29	815	339	1180	1125	385	356
C_3H_7	43	788	312	1180	1110	385	341
C_4H_9	57	777	297	1180	1102	384	332
C_5H_{11}	71	769	287	1180	1098	383	327
C_6H_{13}	85	763	280	1180	1095	382	324
C_7H_{15}	99	760	272	1180	1093	382	320
	∞	730	241	1179	1085	381	306

From table 8 it follows that 2-methyl paraffins should possess another very constant frequency somewhere near 390 cm^{-1} (ν_4) and another fairly constant one in the neighbourhood of 800 cm^{-1} (ν_1). The data in table 7 show that this is indeed the case, ν_1 appearing to vary between 795 and 825 cm^{-1} and ν_4 to lie very close to 435 cm^{-1} . The observed variation in ν_1 along the series is to higher values whereas the predicted variation is to lower values. A similar lack of agreement was found for the corresponding frequency in the 2,2-dimethyl paraffins. No explanation of this can be offered, but possibly a more complete force field using interaction terms would remove both discrepancies. Tentative assignments have also been made from the data for ν_2 and (in a few cases) ν_6 , but these should be treated with reserve until further evidence is available.

iso-BUTANE

The foregoing investigation of the 2-methyl paraffins includes *iso*-butane as the first member of the series. Just as a fuller discussion of neopentane was given to justify the analogous treatment of the 2,2-dimethyl paraffins, so it is useful now to consider *iso*-butane in a little more detail. The first point to note is that alternative assignments of the *iso*-butane skeletal vibrations ν_2 , ν_3 and ν_4 have been proposed

by other workers. Thus Wagner (1939) and Rasmussen (1948) have suggested that n_3 should be assigned to 965 cm.^{-1} , while Pitzer & Kilpatrick (1946) prefer 925 cm.^{-1} . The arguments for and against the present assignment of n_3 to 1170 cm.^{-1} are essentially the same as those concerning the n_3 vibration of neopentane, and need not be repeated. In the case of n_2 and n_4 the alternative assignment consists of interchanging the values given in table 7, i.e. taking $n_2 = 435\text{ cm.}^{-1}$ and $n_4 = 370\text{ cm.}^{-1}$ (Ananthakrishnan 1936; Pitzer & Kilpatrick 1946; Rasmussen 1948).

The present interpretation of the skeletal frequencies of the 2-methyl paraffins provides a scheme in which *iso*-butane appears logically as the first member of the series. The alternative assignment for n_2 and n_4 destroys this consistency. Furthermore, from the relative magnitudes of the moments of inertia of *iso*-butane (Pitzer & Kilpatrick 1946), it would seem that the parallel frequencies of this symmetrical top molecule should show a pronounced Q branch, which would not be prominent for the perpendicular frequencies. Hence the experimental observation of the doublet contour of the 437 cm.^{-1} band (*American Petroleum Institute Research Project 44, Reports 1943-8*), suggests that this is indeed the perpendicular n_4 frequency. It must, however, be admitted that the present interpretation is not in agreement with the available polarization data (Ananthakrishnan 1936).

Finally the applicability of the simple valency force field to *iso*-butane may be checked by an equation due to Lechner (1932) analogous to equations (1) and (2) which were used for neopentane. This equation is

$$\frac{n_1^2 n_2^2}{n_3^2 n_4^2} = \frac{4(3m + M) \cos^2 \phi}{M \left(1 + \cos^2 \phi + \frac{3m \sin^2 \phi}{M} \right)}, \quad (3)$$

where M is the mass of W and m that of X .

Assuming that the angles between the bonds are tetrahedral, the discrepancy between the two sides is 29.5 % using the assignment of table 7. It is important to realize that this discrepancy is quite sensitive to the value of ϕ . If the angles between the bonds are tetrahedral, this implies that $\phi = 70^\circ 32'$. However, electron diffraction measurements (Beach 1940) indicate the higher value $\phi = 72^\circ 39'$. When this is used the discrepancy is reduced to 10.7 %, and a corresponding improvement between calculated and observed values for the frequencies not unnaturally follows; the force constants employed were $k = 4.10 \times 10^5$ dyne/cm. and $k' = 0.46 \times 10^5$ dyne/cm. It is probable that a similar improvement could have been obtained with a modified value of ϕ in the calculations on the series of 2-methyl paraffins, had the relevant experimental data been available for these molecules. It should also be noted that calculations made with equation (3) using the experimental value of ϕ and $n_1 = 795\text{ cm.}^{-1}$ provide a further argument in favour of the present assignment of n_2 , n_3 and n_4 as may be seen from the following figures:

n_2 (cm.^{-1})	n_3 (cm.^{-1})	n_4 (cm.^{-1})	discrepancy (%)
435	1170	370	70
370	965	435	30.9
435	965	370	150
370	915	435	45.6
435	915	370	178

These show that the discrepancies for all the alternative assignments are much larger than that for the present assignment.

GENERAL CONCLUSIONS

Although this investigation started with the limited objective of accounting for certain empirical correlations between the vibration spectra of certain similar branched chain hydrocarbons, the explanation developed has wider implications in the more general question of the interpretation of hydrocarbon spectra, which should be briefly considered. There has been an increasing tendency to attribute all the intense absorption bands in hydrocarbon spectra to vibrations involving hydrogen atoms and to assign the frequencies of the carbon skeleton to the weaker absorption bands. This is not unnatural, since in the spectra of the *n*-paraffins all the intense bands in the region 3 to 16 μ have been shown to be due to CH stretching or deformation vibrations. This presumably arises because the charge on each of the carbon atoms is virtually equal in such hydrocarbons and so the skeletal vibrations which involve stretching of the carbon-carbon bonds are unlikely to give rise to a very effective vibrating dipole. However, when substituents are introduced into the chain, the charge on the carbon atom to which the substituent is attached must be altered considerably and possibly the charges on the near neighbours may also be influenced. Thus it is to be expected that vibrations involving just these parts of the skeleton, where the branching of the chain has occurred, will cause appreciable variations in dipole moment and so become intense in absorption. The fact that the spectra of the highly substituted branched paraffins are much richer and more intense than those of the unsubstituted or singly substituted paraffins thus receives a simple and natural explanation. Otherwise the large number of new bands in the highly methylated paraffins have to be accounted for as arising from various deformation motions of the methyl groups. In view of the constancy of the assigned deformation frequencies of CH_3 and CH_2 groups in all such hydrocarbons this seems rather unlikely.

The second point is that there has been a general tendency to assign all absorption frequencies higher than 1100 cm^{-1} to hydrogen deformation frequencies rather than to skeletal frequencies, thus restricting the spread of the skeletal stretching frequencies to the range 1100 to 750 cm^{-1} . Since the C-C stretching fundamental in ethane is at 990 cm^{-1} , this would seem to imply a reduction in the C-C bond force constant in going from ethane to complex hydrocarbons. On the other hand the present treatment, which in this respect resembles that of Ahonen (1946) gives a range of approximately 1300 to 700 cm^{-1} for these frequencies. This implies no lowering of the bond force constant in going from ethane to the complex hydrocarbons and almost certainly indicates a small increase. Unfortunately there does not appear to be any way of obtaining independent evidence on this point. For instance, if the force constant of the C-C bond varies between one hydrocarbon and another (and possibly between the various C-C bonds in one hydrocarbon) this should be accompanied by small changes in the corresponding bond distances, but these distances have not yet been sufficiently accurately determined to allow such a possible change to be detected.

There is, however, independent evidence that C-C skeletal stretching frequencies can reach values between 1100 and 1300 cm^{-1} , the most direct coming from the vibration spectrum of diamond, which has a Raman frequency at 1332 cm^{-1} and a group of strong infra-red absorption frequencies between 1350 and 1200 cm^{-1} (Robertson, Fox & Martin 1934). Another way of approaching this question is through the use of isotopic atoms, the complete replacement of hydrogen by deuterium separating immediately the hydrogenic from the skeletal frequencies (Sheppard & Sutherland 1947). Unfortunately no deuterated branched hydrocarbon is yet available, but recently some very interesting results have been reported on partially deuterated *n*-propane (McMurray, Thornton & Condon 1948). These indicate that the assignment of 1052 cm^{-1} to a skeletal mode is wrong, since this frequency is strongly modified in 2-deutero propane and make it seem likely that in *n*-propane the C-C skeletal frequency is to be associated with the infra-red band at 1150 cm^{-1} which is not affected by deuteration.

Although the foregoing treatment indicates that for the 2, 2-dimethyl and 2-methyl paraffins a simple valency force field is adequate for the calculation of certain skeletal frequencies, it will be observed that the sets of force constants used for the two series are not identical, and differ also from those used in the discussions of the symmetrical molecules neopentane and *iso*-butane. These differences must arise in part from the fact that the treatment is approximate; in particular, interaction terms and the hydrogen vibrations have been ignored. Moreover the variations in particular frequencies from molecule to molecule within a given series have been attributed to variations in the mass of the molecule. For all these reasons it is not justifiable without further investigation to consider that the above differences in force constants are significant and that they indicate variations in bond strength.

We wish to thank Mr R. H. Merson who made some preliminary calculations for us on these lines which indicated that this method of approach was likely to be profitable, and to express our appreciation of the helpful discussions we have had with Dr N. Sheppard.

REFERENCES

- Ahonen, C. O. 1946 *J. Chem. Phys.* **14**, 625.
American Petroleum Institute, Research Project 44, 1943-8. *Catalogue of infra-red absorption spectral data*.
Ananthakrishnan, R. 1936 *Proc. Ind. Acad. Sci. A*, **3**, 527.
Avery, W. H. & Ellis, C. F. 1942 *J. Chem. Phys.* **10**, 10.
Beach, J. Y. 1940 *J. Chem. Phys.* **8**, 303.
Fellgett, P. B., Harris, G. P., Simpson, D. M., Sutherland, G. B. B. M., Thompson, H. W., Whiffen, D. H. & Willis, H. A. 1946 *Institute of Petroleum Reports*.
Fenske, M. R., Braun, W. G., Wiegand, R. V., Quiggle, D., McCormick, R. H. & Rank, D. H. 1947 *Industr. Engng Chem. (Anal. ed.)*, **19**, 700.
Kohlrausch, K. W. F. & Köppl, F. 1934 *Z. phys. Chem. B*, **26**, 209.
Lechner, F. 1932 *Wien. Ber.* **141**, 633.
Lechner, F. 1933 *Anz. Akad. Wiss. Wien.* **70**, 126.
McMurray, H. L., Thornton, V. & Condon, F. E. 1948 *J. Opt. Sci. Amer.* **38**, 1095.
Pitzer, K. S. & Kilpatrick, J. E. 1946 *Chem. Rev.* **39**, 435.
Rank, D. H. & Bordner, E. R. 1935 *J. Chem. Phys.* **3**, 248.

- Rasmussen, R. S. 1948 *J. Chem. Phys.* **16**, 712.
Robertson, R., Fox, J. J. & Martin, A. E. 1934 *Phil. Trans. A*, **232**, 463.
Rosenthal, J. E. 1934*a*. *Phys. Rev.* **45**, 538.
Rosenthal, J. E. 1934*b* *Phys. Rev.* **46**, 730.
Rosenthal, J. E. 1935 *Phys. Rev.* **47**, 235.
Sheppard, N. 1948 *J. Chem. Phys.* **16**, 690.
Sheppard, N. & Sutherland, G. B. B. M. 1947 *Nature*, **159**, 739.
Sheppard, N. & Sutherland, G. B. B. M. 1949 *Proc. Roy. Soc. A*, **196**, 195.
Silver, S. 1940 *J. Chem. Phys.* **8**, 919.
Simpson, D. M. & Sutherland, G. B. B. M. 1947 *J. Chem. Phys.* **15**, 153.
Slawsky, Z. I. & Dennison, D. M. 1939 *J. Chem. Phys.* **7**, 522.
Straley, J. W., Tindal, C. H. & Nielsen, H. H. 1942 *Phys. Rev.* **62**, 151.
Sutherland, G. B. B. M. & Thompson, H. W. 1943 *Petroleum Board, Enemy Oils and Fuels Committee Report*.
Voge, H. H. & Rosenthal, J. E. 1936 *J. Chem. Phys.* **4**, 137.
Wagner, J. 1939 *Z. phys. Chem. B*, **45**, 341.
Wall, F. T. & Eddy, C. R. 1938 *J. Chem. Phys.* **6**, 107.
Young, C. W., Koehler, J. S. & McKinney, D. S. 1947 *J. Amer. Chem. Soc.* **69**, 1410.
-

The use of polarized infra-red radiation in the study of doubly oriented long-chain polymers

BY E. J. AMBROSE, A. ELLIOTT AND R. B. TEMPLE

Courtaulds, Limited, Maidenhead, Berks

(Communicated by A. H. Wilson, F.R.S.—Received 19 January 1949)

The use of polarized infra-red radiation in examining the structure and orientation of high polymers has been investigated quantitatively. It is shown that infra-red spectroscopy can furnish evidence for double orientation in rolled sheets of nylon 66, polyvinyl alcohol and (with less certainty) polythene. In the case of polyvinyl formate, acetate, chloride and polyvinylidene chloride such double orientation could not be detected.

Evidence is given to show that in nylon the N-H bond is bent by hydrogen bonding forces, the angle between this bond and the plane of the skeleton being thereby reduced from 39° (the valency angle) to 22°.

The structure of polyvinyl alcohol recently proposed by Bunn receives strong support from the absence of dichroism in the O-H frequency in the spectrum of that material, when a doubly oriented specimen is examined.

INTRODUCTION

This paper contains the results of a re-examination, under higher resolving power, of the polymers dealt with in an earlier paper (Elliott, Ambrose & Temple 1948). Observations are confined to the spectral region between 2500 and 4000 cm.⁻¹, using a lithium fluoride prism, and in carrying out this work our object has been to place the method on a more quantitative basis than hitherto.

In spite of the fact that in certain respects long-chain polymers are easier to deal with than simple crystals, the amount of work already published on such materials in the fundamental region from 2.5 to 14 μ is small and is mostly confined to brief references to work still to be published (Thompson & Mann 1948; Wright 1948).

The observations of Sutherland & Vallance-Jones (1947) on the polythene bands at 6.9, 7.3 and 13.9μ have been confirmed and extended by us (Elliott *et al.* 1948); some of our results are also in general agreement with those of Glatt & Ellis (1947, 1948) on oriented polythene and nylon in the overtone region, although interesting and so far unexplained discrepancies also occur. These are discussed below.

EXPERIMENTAL

It is extremely difficult to prepare uniformly very thin and highly oriented polymer films of sufficient size (15×4 mm.) for use with our existing apparatus, but with the technique now adopted such variations of thickness have little effect on the results.

As already mentioned (Elliott *et al.* 1948), polymer films may be oriented either by stretching or rolling, the choice of the method used depending on the material. For example, the vinyl polymers orient easily on stretching, whereas polythene and nylon cast from solution stretch only with difficulty; when suitably protected, however, they are readily rolled out. Polythene is most suitably protected by being melted between two pieces of silver chloride sheet before rolling; the silver chloride, being transparent to 12μ , need not be subsequently removed, although its high refractive index causes an appreciable loss of energy. It can be readily removed from all but the thinnest films by soaking in sodium thiosulphate solution. A different technique is necessary with nylon, since this reacts at the melting-point with silver chloride. The nylon is cast from solution on to a cadmium sheet, melted into contact with the metal and, after rolling, detached by soaking in water. A few polymers (e.g. polyvinyl alcohol) respond to rolling only at high temperatures, the rolls being suitably pre-heated by an air blast.

According to X-ray evidence (Bunn 1939; Bunn & Garner 1947) double orientation of the polythene and nylon crystallites is produced on rolling, the chains being oriented in the direction of rolling and, in addition, certain crystalline planes being arranged parallel to the polymer sheet. The orientation, however, differs from that of a true crystal since although the chains are parallel they need not be in the same sense but may be rotated through 180° with respect to each other. This type of orientation can be detected by polarized infra-red radiation arranged to traverse the material in a direction as nearly parallel to the polymer chains as possible.

The spectrometer arrangement differs from that previously described (Elliott *et al.* 1948) only in the substitution of a lithium fluoride prism which at 3μ results in a sevenfold gain in dispersion over that formerly possible. The slit width usually employed was 0.085 mm., and the equivalent slit width of 8 cm.^{-1} is small compared with the widths of many of the bands observed, although some of the polythene bands are so narrow that still greater resolution would be desirable. The present resolution is limited by the detector-sensitivity.

To examine a specimen for double orientation the arrangement of prisms shown in exploded form in figure 1 is used. These prisms are made of potassium bromide ($\mu \approx 1.53$) to match approximately the polymers examined. The film is placed between the hypotenuse faces of the prisms which are contained in a brass cell cemented to

plane windows. For the near infra-red region it is convenient to use silica windows 0.5 mm. thick, since two such windows are sufficiently transparent in the region of the hydrogen stretching frequencies. To make optical contact, a small amount of immersion liquid is placed in the cell. Very few suitable liquids are available in this region, the most generally useful being a solution of carbon tetrabromide (18 g.) in carbon tetrachloride (10 ml.), the refractive index of this mixture being 1.53. It has the advantage of being partially self-sealing, slight evaporation of the volatile constituent leaving a seal of solid CBr_4 .

Carbon disulphide and tetrachloroethylene have also been employed occasionally. When using carbon disulphide the cell windows are luted with a fish glue.

The cell containing the specimen is mounted on a carriage which is moved to allow the radiation to pass alternately through the specimen and through the clear part of the prisms. Two specimens of suitable size are cut, with the polymer chain axes respectively parallel and perpendicular, to the length of the strip. These specimens are mounted in turn between the prisms as shown in figure 1. For each specimen, records are made with the electric vector of the radiation parallel and perpendicular, respectively, to the spectrometer slit. The deductions which can be made are described later.

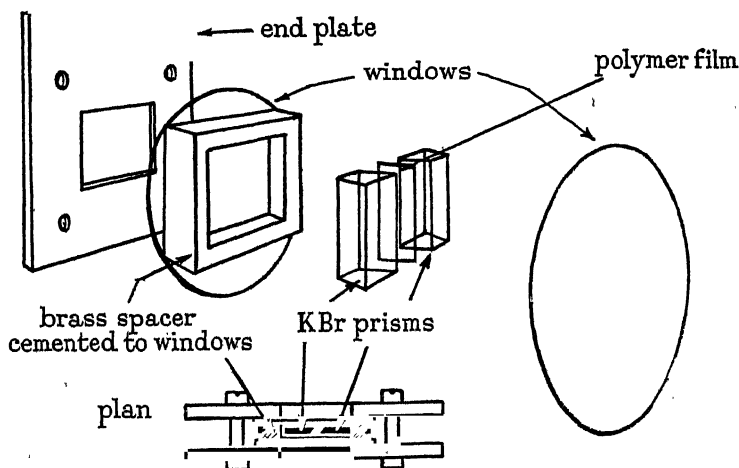


FIGURE 1

THEORETICAL CONSIDERATIONS

Consider a right-handed system of axes $Oxyz$ as in figure 2, with a beam of polarized radiation incident in the direction Ox . A doubly oriented polymer sheet occupies the plane AOz with the chain axis parallel to Oz , and OP is the direction along which a change in dipole moment of the polymer molecules will give rise to absorption of radiation. OP may be defined by the polar angle θ relative to the axis z and the azimuth angle ϕ relative to the plane xOz . The plane zOA has azimuth α .

It is intended to calculate the energy absorbed when the electric vector of the radiation is parallel to Oy and Oz respectively. Although this energy is required in terms of θ , ϕ and α , the calculation is simplified by using a different co-ordinate

system in which OB (normal to the polymer sheet) is the polar axis and the plane zOB is the azimuthal plane. The direction OP is now defined by the polar angle ψ and the azimuth angle χ (see figure 2).

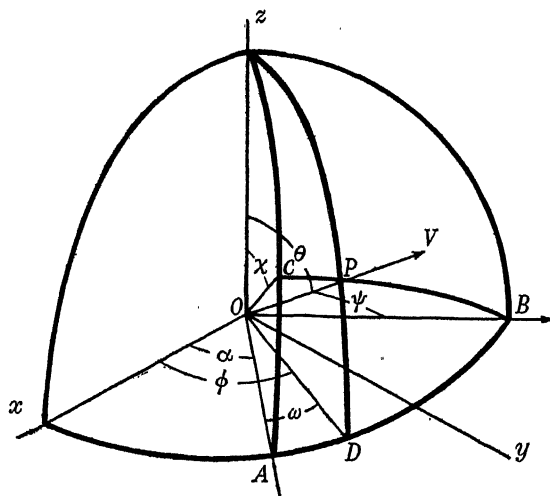


FIGURE 2

Taking a vector V along OP , this may be resolved into the components

$$V(\sin \psi \sin \chi \sin \alpha + \cos \psi \cos \alpha), \quad (1)$$

$$V(\sin \psi \cos \chi), \quad (2)$$

along Oy and Oz respectively.

In a completely doubly oriented crystalline polymer, although the crystallites will have their chain axes all parallel to Oz and crystallographic planes parallel to the sheet OAz , there will be equal numbers lying in the two opposite directions along z and right- and left-handed, as it were, with respect to the plane OAz . We must therefore consider vectors obtained by reflecting OP in the planes OAz and xOy . The co-ordinates of all the vectors are (ψ, χ) , $(\pi - \psi, \chi)$, $(\psi, \pi - \chi)$, $(\pi - \psi, \pi - \chi)$. Four more vectors could be obtained by reflecting each of these in the plane zOB , but these will be in the same direction as the first set, though of opposite sign. The vector V represents the change of dipole moment for the corresponding normal vibration of a molecule near the equilibrium position, and the radiation absorbed per molecule is proportional to V^2 . This is independent of the sense of the vector, hence we need only consider the four vectors specified above. The radiation absorbed is proportional to the square of the scalar product of the electric vector E and the vector V . If there are N molecules in all, the energy absorbed by $\frac{1}{4}N$ molecules with dipole moment changes parallel to OP and E vectors parallel to Oy is

$$\frac{1}{4}KN(\sin \psi \sin \chi \sin \alpha + \cos \psi \cos \alpha)^2.$$

By adding the contribution from the molecules with moment change parallel to the other three vectors, the energy absorbed with E vector parallel to Oy is

$$KN(\sin^2 \psi \sin^2 \chi \sin^2 \alpha + \cos^2 \psi \cos^2 \alpha).$$

The contributions of all the molecules when the E vector is parallel to Oz are identical. Therefore the energy absorbed with E vector parallel to Oz is

$$KN(\sin^2 \psi \cos^2 \chi).$$

The dichroic ratio (which can be measured experimentally), taken as the ratio of optical density with the E vector parallel to Oz to the corresponding quantity with the E vector parallel to Oy , is

$$\frac{\sin^2 \psi \cos^2 \chi}{\sin^2 \psi \sin^2 \chi \sin^2 \alpha + \cos^2 \psi \cos^2 \alpha} = A. \quad (3)$$

The position of the polymer with the chains parallel to Oz is called the A position. We now consider the case where the specimen is rotated about OB (figure 2) through a positive angle $\frac{1}{2}\pi$ to bring the chain's axis parallel to OA (B position). The angle ψ is not affected by this operation, and χ is changed to $\chi + \frac{1}{2}\pi$. The value of the dichroic ratio (defined as above) is now

$$\frac{\sin^2 \psi \sin^2 \chi}{\sin^2 \psi \cos^2 \chi \sin^2 \alpha + \cos^2 \psi \cos^2 \alpha} = B. \quad (4)$$

The direction of the vector parallel to OP could be calculated from (3) and (4) in terms of ψ and χ , but it is more convenient to have the vector referred to the chain axis as polar axis and the plane of the polymer sheet as azimuth plane. This requires the direction to be defined in terms of θ and $\phi - \alpha = \omega$.

By resolving a unit vector in the direction OP along the directions OA , OB and Oz respectively it is found that

$$\sin \theta \cos \omega = \sin \psi \sin \chi, \quad (5)$$

$$\sin \theta \sin \omega = \cos \psi, \quad (6)$$

$$\cos \theta = \sin \psi \cos \chi. \quad (7)$$

Making use of these relations we find

$$A = \frac{\cos^2 \theta}{\sin^2 \theta \cos^2 \omega \sin^2 \alpha + \sin^2 \theta \sin^2 \omega \cos^2 \alpha}, \quad (8)$$

$$B = \frac{\sin^2 \theta \cos^2 \omega}{\cos^2 \theta \sin^2 \alpha + \sin^2 \theta \sin^2 \omega \cos^2 \alpha}. \quad (9)$$

Solving (8) and (9) for θ and ω we have

$$\cot^2 \theta = \frac{A(\cos^2 \alpha + B \sin^2 \alpha \cos^2 \alpha)}{1 + B \cos^2 \alpha + AB \sin^2 \alpha (1 - 2 \sin^2 \alpha)}, \quad (10)$$

$$\sin^2 \omega = \frac{1 - AB \sin^4 \alpha}{1 + B \cos^2 \alpha + AB \sin^2 \alpha (1 - 2 \sin^2 \alpha)}. \quad (11)$$

These formulae have been worked out on the assumption that there is complete linear and planar orientation of the crystallites in the polymer, that is, complete double orientation in the sense in which the expression is used by Bunn & Garner (1947).

When the orientation of the crystallites is incomplete, we can think of effective values of the co-ordinates $\bar{\theta}$ and $\bar{\omega}$ such that substitution of these values in expressions (10) and (11) would give the observed values of the dichroic ratios A and B . These effective values will depend on the direction which the valency bond makes with the molecular axes, and also on the degree of orientation in the specimen. Although the connexion between the effective values $\bar{\theta}$ and $\bar{\omega}$ and the angles which define the orientation of the crystallites themselves is not yet clear, a good deal can be deduced from internal comparisons for different modes of vibration.

For a completely unoriented polymer, $A = B = 1$. Substituting these values in equations (10) and (11) we find $\bar{\omega} = 45^\circ$, $\bar{\theta} = 54^\circ 42'$. The angle $\bar{\omega}$ will always be 45° when there is no planar orientation, and $\bar{\theta}$ will be 0° when the dipole moment change takes place only along the direction of stretching or rolling (in this case, of course, the angle ω has no meaning). For a change of dipole moment perpendicular to the direction of rolling, $\bar{\theta}$ has the value 90° .

It will be noticed that the conception used here differs from that in our earlier paper. In our present work, we have taken the more general case where the dipole moment can take up any position relative to the orientation axis, and do not consider specifically the orientation of the axis itself.

CORRECTION FOR THE FINITE ANGLE OF THE CONE OF INCIDENT RADIATION

In the derivation described above it has been assumed that the radiation is incident in the unique direction Ox . In practice it is necessary to use a converging beam of radiation, and in the Perkin-Elmer spectrometer the cone of radiation subtends a semi-angle of 5° . If necessary a correction can be applied, but in all cases referred to in this paper, it is negligible.

TABLE 1. DICHOIC RATIOS AND ORIENTATION ANGLES

polymer	band	wave no. (cm. ⁻¹)	$\alpha = 36^\circ$ dichroic ratio		orientation angles	
					$\bar{\theta}$	$\bar{\omega}$
			A	B		
nylon 66	N-H	3308	0.63	1.73	61.6	39.1
	?	3078	0.72	2.11	60.7	34.6
	CH ₂ as.	2941	0.44	0.71	63.5	53.4
	CH ₂ s.	2866	0.68	1.54	60.7	40.1
polyvinyl	CH ₂ as.	2945	0.35	0.82	67	51.5
alcohol	CH ₂ s.	2910	0.35	1.38	67.5	44.5
polythene	CH ₂ as.	2926	0.51 ^s	1.05	63	47.4
	CH ₂ s.	2855	0.42 ^s	1.14	65.2	46.4

OBSERVATIONS ON ORIENTED POLYMERS

Nylon 66

Nylon 66 was prepared by the method described by Chambret (1947), in which crystals of hexamethylene di-ammonium adipate are condensed to a polymer in the solid phase. The condensation takes place at 180°C in a moderate vacuum. The product is highly crystalline and gives a very sharp X-ray pattern. Rolled specimens

were examined in the *A* and *B* positions as described above. The records are shown in figure 3 *a* and *b* respectively. As expected, in the *A* position, the band at 3308 cm.^{-1} , owing to the NH stretching vibration, shows stronger absorption perpendicular to the chains, because the NH bond lies in a plane perpendicular to the chain axis. In the *B* position, where one is looking along the chains, the absorption is stronger with the electric vector parallel to the sheet, indicating that the NH bond is bent by the hydrogen bonding forces into the plane of the sheet.

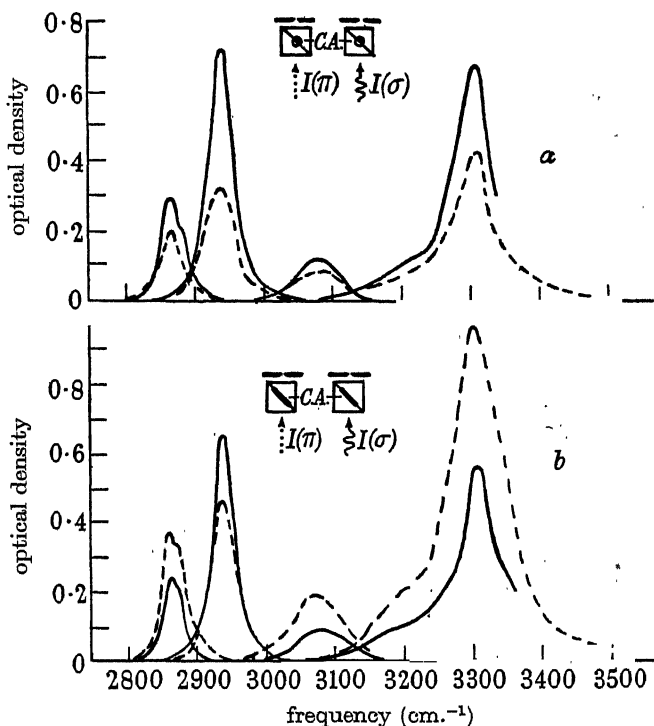


FIGURE 3. *CA*, carbon chain axis.

The symmetric methylene frequency shows the same type of orientation as the NH frequency whereas the anti-symmetric methylene frequency shows stronger absorption perpendicular to the sheet. This is owing to the fact that in the symmetrical CH_2 stretching vibration the dipole moment changes along the axis of the group, whereas in the anti-symmetrical mode it changes at right angles to the axis.

The values of the dichroic ratios *A* and *B* are given in table 1, along with the values of $\bar{\theta}$ and $\bar{\omega}$ calculated from equations (10) and (11). The value of $\bar{\omega}$ for the symmetric frequency of the CH_2 group is 40.1° , and in this case the dipole moment change is collinear with the axis of the CH_2 group. For the anti-symmetric frequency, $\bar{\omega}$ is 53.4° , and since in this case the dipole moment change is at right angles to the CH_2 axis, the effective direction of this axis deduced from the anti-symmetric frequency is $90^\circ - 53.4^\circ = 36.6^\circ$. It appears justifiable to ascribe the difference between the two values to experimental error, and to take the mean value of 38.3° as the effective direction of the CH_2 axis. This value is almost identical with $\bar{\omega}$ for the NH bond (39.1°).

It is convenient to use the stereographic projection invented by Ptolemy to represent the three-dimensional directions of the dipole moment change, electric vector of the incident radiation, etc. It is adequate for our purpose, since only directions, not distances, as such are measured.

A complete description of the properties of the stereographic projection has been given by Sohon (1941), but it may be noted here that directions are represented by lines passing through the centre of a sphere and terminating on the surface, the points where these lines terminate being represented by projection on to the equatorial plane of the sphere, by drawing a line from the south pole of the sphere to the point in question. In the case of a point in the southern hemisphere, it is projected by joining the point on the sphere to the north pole. Planes passing through the centre of the sphere intersect its surface in great circles. These great circles project on to the equatorial plane as true circles.

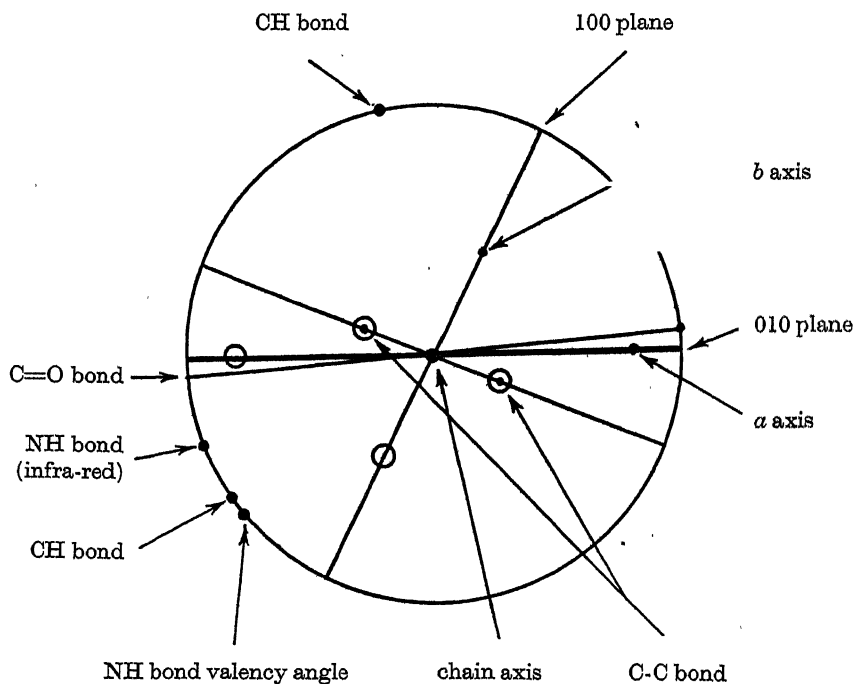


FIGURE 4

Figure 4 shows a stereographic representation of the structure of nylon 66 as determined by Bunn & Garner. Crystallographic planes parallel to the chain axis are represented by great circles through the poles and the direction of various valency bonds and the changes of dipole moment associated with molecular vibrations are represented by points on the stereogram. This method of representation enables one to compare any proposed structure with the results obtained from the polarized infra-red absorption measurements.

In nylon 66 the axial and planar orientation of the hydrogen atoms is incomplete and it is not possible to compare the valency angles with those determined by Bunn

& Garner directly; it is possible, however, to make internal comparisons. The angles which the bonds in nylon 66 make with each other and with the crystallographic planes have been read from the diagram given by Bunn & Garner. In this way we find that the angle between the 010 plane and the plane of the carbon chain is about 22° . The plane of the carbon chain contains the vector V for the symmetric methylene stretching vibration and is perpendicular to V for the anti-symmetric vibration. The inclination of the axis of the CH_2 group to the plane of the sheet (derived from polarized infra-red measurements) coincides with that of the NH stretching frequency and therefore the NH bond is inclined at an angle of 22° to the 010 plane. The sign of this angle will now be considered. Because of the possible inversion of the nitrogen valencies, the hydrogen atom could lie on either side of the carbon chain if there were no hydrogen bonding. It seems reasonable to choose the position which brings the NH bond towards the direction of the $\text{C}=\text{O}$ bond, which makes an angle of about 4° to the 010 plane. In this case the angle between the NH bond and the $\text{C}=\text{O}$ bond is $22^\circ - 4^\circ = 18^\circ$. If no hydrogen bonding forces were operating, the nitrogen atom would lie in the plane of the carbon chain and the NH bond would be inclined to this plane at an angle corresponding to the nitrogen valency angle. In this case the NH bond would be inclined at an angle of 39° to the 010 plane. The hydrogen bonding force has reduced this angle from 39° to 22° . Whether this is due to a distortion of the NH bond or a distortion of the carbon-nitrogen-carbon chain is unknown. Alternatively, the N-H bond could lie in the same plane as the carbon-carbon chain. This would require a planar configuration for the nitrogen valencies such as has been proposed by Corey for diketopiperazine (Corey 1938). The distance between the hydrogen and oxygen atoms forming the hydrogen bond would, in this case, be $2.0 \text{ \AA}$ as compared with $1.9 \text{ \AA}$ in the first arrangement. With this position for the hydrogen atom it is somewhat difficult to understand why the $\text{C}=\text{O}$ bond needs to be distorted away from the plane of the carbon-carbon chain, and why all the atoms involved in hydrogen bonding do not lie in this plane. Further information on this point should be forthcoming when the doubly oriented specimens are investigated in the region of the $\text{C}=\text{O}$ and CN frequencies. It may also be necessary to consider different orientation of the amide and hydrocarbon portions of the chain.

The origin of the band at 3078 cm.^{-1} is not known. It occurs as a characteristic band in polypeptides (Darmon & Sutherland 1947) and may be an overtone. More recently, Astbury, Dalgleish, Darmon & Sutherland (1948) have attributed it to a new type of hydrogen bonding. The change of moment associated with this vibration subtends a smaller angle than the NH bond to the 010 plane.

Glatt & Ellis (1948) have studied the spectra of nylon 6-10 in the overtone region, using polarized radiation. They have found, in agreement with our results for nylon 66, that the dipole moment change of the NH deformation frequency takes place in a direction parallel to the chains. They also find that the third harmonic of the $\text{C}=\text{O}$ vibration vanishes completely in the parallel position; this is unexpected, since we previously found only moderate dichroism in the fundamental $\text{C}=\text{O}$ vibration at about 1650 cm.^{-1} in nylon 66.

Polyvinyl alcohol

Earlier work (Elliott *et al.* 1948) on hot-stretched polyvinyl alcohol with a rock-salt prism revealed only low values for the dichroism of the CH stretching frequencies, while no dichroism was found in the OH band. Since this work was done, a new structure for the crystallites in polyvinyl alcohol has been proposed by Bunn (1948), based on X-ray intensity measurements. This structure suggests that polyvinyl alcohol should show double orientation when rolled, though probably to a less degree than nylon 66. Attempts to orient the material by cold rolling were unsuccessful. A twofold extension could be achieved by the use of great force, but the birefringence shown was extremely low; further experiments on the polymer dyed with Congo red confirmed that this was due to very low orientation. Congo red, applied to oriented polymers, shows high dichroism in the visible region and is a convenient indicator of orientation (Preston 1933; Land & West 1946). It was found that the cold-rolled film showed practically no visible dichroism in spite of a twofold extension.

Hot rolling at about 50° C was found to produce satisfactorily oriented films, using the technique described above.

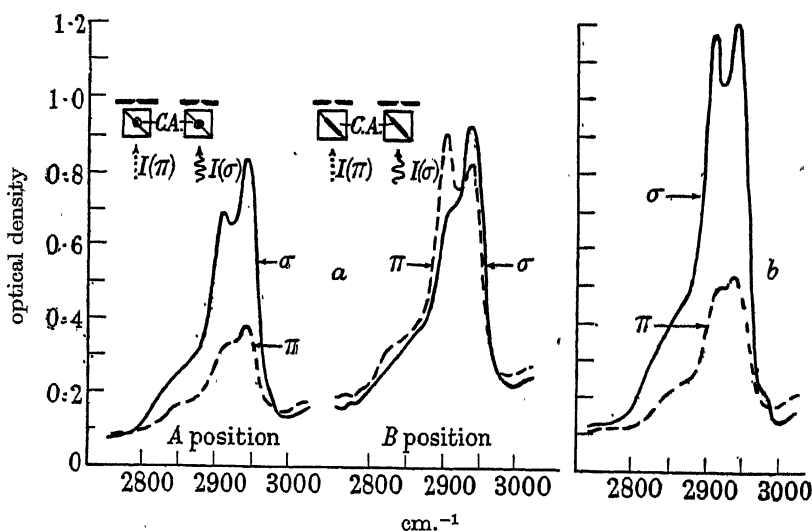


FIGURE 5. The spectra of polyvinyl alcohol, *a*, hot rolled 5 × and *b* hot stretched. CA, carbon chain axis.

The spectrum of polyvinyl alcohol contains a very broad, strong band centred at 3345 cm.⁻¹, attributed to the OH stretching frequency. In our earlier work we stated that the *breadth* of the OH band could not easily be recognized with an exact crystalline structure such as proposed by Mooney. Mr Glatt has, however, pointed out to us (private communication) that hydrogen bonding produces a broad band even in a highly crystalline material (for instance, succinic acid). In addition to the OH band there are two strong partially resolved bands, at 2910 and 2945 cm.⁻¹, and a much weaker one at about 2865 cm.⁻¹. These spectra are shown in the *A* and *B* positions in figure 5*a*.

The partially resolved strong bands are attributed to the symmetrical and anti-symmetrical modes of vibration of the CH_2 groups respectively. The weak band may be due to the single CH bond attached to the same carbon atom as the O atom. The orientation in the direction of rolling is quite marked, the dichroic ratio D_π/D_σ for the *A* position for the CH_2 frequencies being 0.35. The presence of double orientation, with certain planes in the crystalline polymer parallel to the surface of the sheet, is shown by the dichroism in the *B* position. Here the symmetrical mode of vibration of the CH_2 group (2910 cm.^{-1}) is stronger in the π position, that is when the electric vector is parallel to the plane of the sheet. The anti-symmetrical mode, on the other hand, is most strongly excited in the σ position (electric vector perpendicular to sheet). This is similar to the results obtained with nylon 66.

No dichroism has been found in the OH band in either the *A* or *B* position. This at once discriminates in favour of Bunn's views on the structure of polyvinyl alcohol, for on Mooney's structure (Mooney 1941), the OH and CH bonds should exhibit about the same degree of dichroism in the *B* position, whereas the random arrangement proposed by Bunn is quite compatible with the absence of dichroism.

The effective directions of dipole moment change have been calculated from equations (10) and (11) for the symmetrical and anti-symmetrical CH_2 frequencies (see table 1). Using the same argument as set out for nylon 66, it is seen that the effective direction of the CH_2 axis obtained from the anti-symmetrical vibration is $90^\circ - 51.5^\circ = 38.5^\circ$. This does not agree as well with the value of $\bar{\omega}$ for the symmetric frequency (44.5°) as in the case of nylon. The lack of agreement does not appear to be due to random errors, as similar results have been obtained with other specimens. It may be caused by an overlapping band (possibly an overtone). This is to some extent supported by the result shown in figure 5*b* which shows the spectrum of a stretched (not rolled) specimen, normal to the incident radiation. The two CH_2 bands show the same type of dichroism here, but the contours in the π -position are not exactly a reduced copy of those in the σ -position, as would be expected in a specimen which, one presumes, has only linear orientation. An overlapping band may also be falsifying the measurement here.

The weak band which is probably due to the single CH bond shows perpendicular dichroism in the *A* position. In the *B* position there is practically no dichroism. This is to be expected, as the CH bond may be on either side of the carbon chain. Possibly weak dichroism could be detected here if there were not overlapping from the much stronger CH_2 frequencies.

Polythene

Dichroism in stretched polythene film has been observed by Sutherland & Vallance-Jones (1947) who observed perpendicular dichroism in several of the CH_2 frequencies. They mentioned that the band at 1375 cm.^{-1} , generally attributed to CH_3 groups, was anomalous. We have observed (Elliott *et al.* 1948) that the 1375 cm.^{-1} band shows parallel dichroism, and suggested that the CH_3 groups may terminate side chains which are long enough to lie parallel to the main chains.

It was hoped that work with higher dispersion in the region of the CH_2 and CH_3 stretching frequencies might throw more light on the question, and also that some

information might be obtained from measurements on doubly oriented polythene film.

Rolled specimens have been examined for dichroism with the radiation incident normally on the surface; the spectra are given in figure 6*a* and *b*. It has generally been found that the highest dichroism of the CH_2 bands occurs with an eightfold linear extension of the polythene. This produces a dichroic ratio of about 4:1, with absorption strongest in the perpendicular position (E vector perpendicular to chain axis). Further extension by rolling produces a considerable reduction in dichroism. For example, an 18-fold overall extension in length produces a dichroic ratio of about 2:1 as shown in figure 6*a*. This treatment causes the polythene to rupture, and holes begin to appear in the film. Because the polythene is embedded in silver chloride, it is difficult to ascertain what the extension really amounts to in such a case. There is some evidence that the bands broaden slightly. Both effects may perhaps arise from distortion of the hydrogen atoms from their normal positions in the crystal lattice, owing to the great mechanical stress applied.

The specimens of polythene whose spectra are shown in figure 6 are about 1μ thick. As a result of the greatly increased resolving power (which enables the true contours of the bands to be observed) much thinner specimens are needed than in our earlier work with a rock salt prism.

The polythene spectra shown in figure 6 contains the symmetrical and anti-symmetrical CH_2 stretching frequencies, at 2855 and 2926 cm^{-1} respectively. These figures are in quite good agreement with those of 2853 and 2925 cm^{-1} obtained by Fox & Martin (1940) from grating spectrometer measurements. The graphical analysis given in figure 6 reveals also two weak perpendicular bands at 2895 and 2965 cm^{-1} . The second band is the anti-symmetrical methyl group frequency referred to by Fox & Martin, from which they inferred the presence of methyl groups on side chains. The perpendicular character of this band supports the view that the methyl groups are on side chains oriented parallel to the main chains. The lower frequency may be due to a CH bond at the carbon atom on the main chain where the side chain is joined. Fox & Martin have suggested 2890 cm^{-1} for the frequency of a $\geq \text{CH}$ group, which is fairly near the value which we find. The perpendicular character of the two weak bands is in harmony with these assignments.

The symmetrical CH_2 frequency should show parallel dichroism. It is to be expected that it would be weaker than the anti-symmetrical mode, and of lower frequency. No such band has been identified, however, and it may be covered by the much stronger CH_2 band at 2855 cm^{-1} .

Glatt & Ellis (1947), working in the overtone region, have found that the symmetrical CH_2 frequency in stretched polythene showed less dichroism than the anti-symmetrical frequency. This is not in agreement with our results, as may be seen from figure 6*a*. With normal incidence, the symmetrical frequency shows somewhat higher dichroism.

Rolled specimens of polythene have been examined for double orientation by the method already described, see figure 6*c*. It was found necessary to leave the polythene between silver chloride to protect it from the solvent action of the immersion medium. Further protection was afforded by sealing the edges with seccotine.

A thin sheet of polythene was extracted with carbon tetrachloride to remove as much as possible of the low molecular weight fraction, before being rolled between silver chloride sheets. Carbon disulphide was used as the immersion medium.

The considerable increase in absorption going from the *A* position (direction of rolling parallel to spectrometer slit) to the *B* position (direction of rolling perpendicular to slit) may be noted. The dichroic ratios are given in table 1.

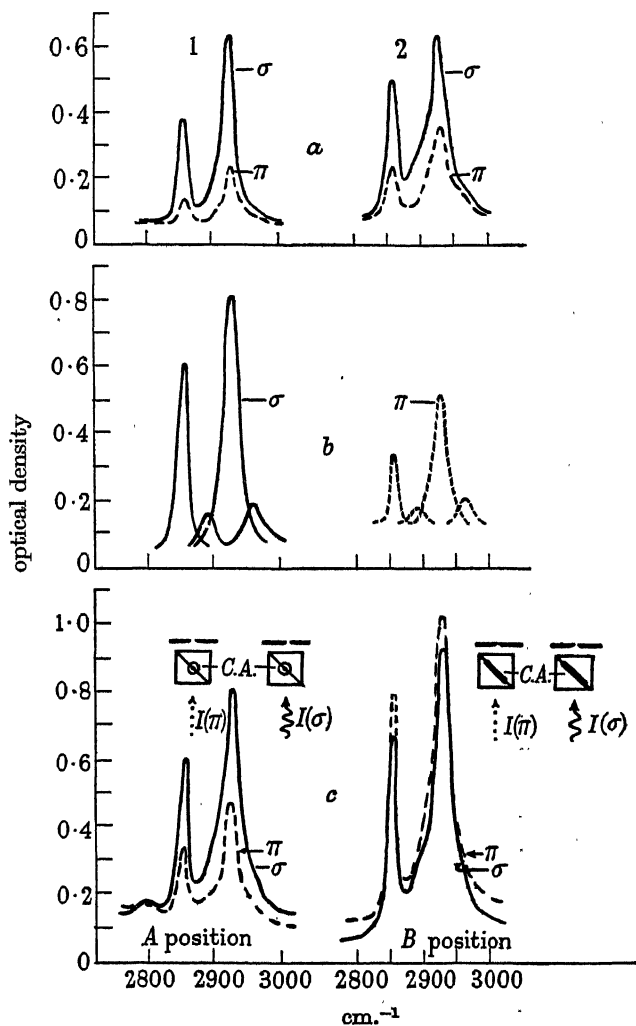


FIGURE 6. The spectra of polythene. *a* Plane of sheet normal to radiation: 1, cold rolled 8 ×, 2, cold rolled 18 ×. *b*, Cold rolled, 8 × (graphical analysis). *c* Cold rolled 8 ×, *A* and *B* positions. *CA*, carbon chain axis.

From the structure for the unit cell of polythene given by Bunn (1939) it is possible to calculate an effective direction of dipole moment change for a completely oriented crystalline polythene sample. The values of $\bar{\omega}$ so obtained are 46.4 and 43.5° for the antisymmetrical and symmetrical frequencies respectively. These effective angles

are near the value $\omega = 45^\circ$ for an unoriented polymer because the planes of the carbon chains of the two molecules in the unit cell are nearly at right angles to one another. Our results show a discrepancy, for an actual imperfectly oriented specimen of polythene should have values of ω between 45 and 46.4° , and 45 and 43.5° , for anti-symmetrical and symmetrical bands, respectively. The actual values found lie outside the calculated range. We have no satisfactory explanation of this discrepancy, but two considerations may be noted:

(a) There is reason for believing that the rolling process produces a strained crystal lattice. This is shown by the fact that excessive rolling lowers the dichroic ratios in the CH_2 bands.

(b) The infra-red spectrum contains a contribution from the amorphous as well as from the crystalline region. The X-ray pattern, on the other hand, shows almost entirely the contribution of the crystalline regions.

Vinyl polymers

We have also examined the following polymers: polyvinyl acetate, kindly given by Shawinigan Electrochemicals Ltd; polyvinyl formate, prepared by our colleague, Mr W. E. Hanby; polyvinyl chloride, an unplasticized specimen of unknown history; polyvinylidene chloride, kindly given to us by the Distillers Co., Epsom. The first two are well known to be completely amorphous, the third is partially crystallizable with great difficulty and the fourth is highly crystalline; unfortunately this last material is not easily prepared in large clear films. None of the spectra has previously been reported under the high dispersion available with the lithium fluoride prism. The spectra are shown in figure 7. All the bands certainly due to the polymers show

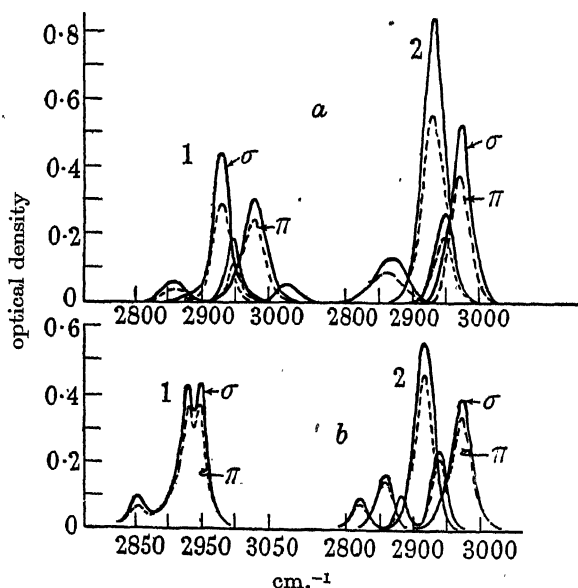


FIGURE 7. *a.* Graphical analyses of the spectra of polyvinyl acetate 18μ thick (1) and polyvinyl formate, 20μ thick (2). *b.* Spectra of polyvinylidene chloride (1) and polyvinyl chloride (graphical analysis) (2).

perpendicular dichroism when examined with the plane of the stretched material perpendicular to the plane of the incident radiation and the dichroic ratios observed are in all cases small, of the order of 1.2 to 1.3.

There are certain features of the spectra which have not yet been cleared up, although it is hoped that further investigation will do so. As already stated, the assignment of the CH_2 stretching frequencies in the spectrum of polythene is unequivocal, the frequencies of 2850 and 2930 cm^{-1} being allocated respectively to the symmetrical and anti-symmetrical vibrations, the latter being the stronger of the two. The spectra of the four polymers named above also show bands in these regions (the former being generally relatively much weaker than in polythene) and this led us to the conclusion that whereas the frequencies remained practically constant from one polymer to the next, the intensities were subject to rather erratic variation. The examination of the rolled polyvinyl alcohol placed between potassium bromide prisms (already discussed above in detail) showed unambiguously, however, that this is incorrect and that the symmetrical CH_2 vibration of this molecule has undergone a shift to appreciably higher frequencies. The weak band also present at about 2860 cm^{-1} in the spectrum of polyvinyl alcohol cannot therefore correspond to the band at the same frequency in polythene and it has been suggested that this is due to the 'lone' hydrogen atoms along the main chain. Polarization measurements with the amorphous or nearly amorphous polymers do not allow us to assign the bands with any certainty and the matter is still further complicated by examination of polyvinylidene chloride. The spectrum of the pure unplasticized material has not been reported before. Thompson & Torkington (1945) investigated the spectrum of 'Saran' between 6 and 20 μ . This is a co-polymer with polyvinyl chloride used technically because of its greater mechanical tractability; its spectrum may contain features not common to either of the pure polymerized constituents. The spectrum of the pure polymer (figure 7*b*, 1) shows two sharp bands, at 2850 and 2940 cm^{-1} respectively, the latter being found to be double when examined under high dispersion, with heads at 2930 and 2947 cm^{-1} . Though this first appears to be a case of accidental degeneracy, a study of the long wave Saran spectrum, published by Thompson & Torkington (1945), does not reveal any simple overtone or binary combination tone which might be responsible, and in spite of the closeness of the two 'components' it is therefore possible that these are really due to the two CH_2 modes, the symmetrical vibration frequency being increased as in polyvinyl alcohol. Examination of the material between the KBr prisms which might at once decide between the two possibilities has been so far prevented by practical difficulties.

The spectra of polyvinyl acetate and chloride are strikingly similar with the exception of two weak bands on the long wave side. (The apparently unoriented one at 2880 is most likely due to a trace of tenaciously held solvent.) These spectra show five main bands; in polyvinyl formate they overlap to such an extent that only a single broad band with two rather indistinct shoulders is found. A graphical separation is roughly possible on the basis of four frequencies but an additional one may be present at about 2890 cm^{-1} . From the similarity of the spectra one is forced to the conclusion that the pattern of bands 2860 cm^{-1} (weak), 2930 cm^{-1} (strong), 2945 cm^{-1} (weak), 2970 cm^{-1} (medium strong) is due to the skeletal unit

$\text{—CH}_2\text{—}\underset{\text{X}}{\text{CH—}}$ and that the nature of the group X has little influence except in the case of the strongly polar —OH group. What is difficult to understand is that even with the acetate group present, where the possibility of two further fundamentals arises, this pattern is not significantly departed from. One must therefore assume that the methyl group frequencies, which in hydrocarbons are to be found at 2962 and 2872 cm.^{-1} according to Fox & Martin (1940), are sufficiently shifted in the polyester to coincide with some of the skeleton frequencies.

An attempt was made to detect planar orientation of the chains in polyvinyl chloride, using the hot-rolled material. Unfortunately, no definite indication of a change of dichroism could be obtained which would have permitted the identification of the CH_2 stretching modes. While insufficient orientation, or penetration of the polymer film by the immersion medium may have been responsible, it is also possible that the real reason is that in these amorphous or poorly crystalline polymers, the intermolecular forces are insufficient to produce such orientation.

REFERENCES

- Astbury, W. T., Dalglish, C. E., Darmon, S. E. & Sutherland, G. B. B. M. 1948 *Nature*, **162**, 596.
- Bunn, C. W. 1939 *Trans. Faraday Soc.* **35**, 482.
- Bunn, C. W. 1948 *Nature*, **161**, 929.
- Bunn, C. W. & Garner, E. V. 1947 *Proc. Roy. Soc. A*, **189**, 39.
- Chambret, F. 1947 *Bull. Soc. chim.* **14**, 283.
- Corey, R. B. 1938 *J. Amer. Chem. Soc.* **60**, 1598.
- Darmon, S. E. & Sutherland, G. B. B. M. 1947 *J. Amer. Chem. Soc.* **69**, 2074.
- Elliott, A., Ambrose, E. J. & Temple, R. B. 1948 *J. Chem. Phys.* **16**, 877.
- Fox, J. J. & Martin, A. E. 1940 *Proc. Roy. Soc. A*, **157**, 961.
- Glatt, L. & Ellis, J. W. 1947 *J. Chem. Phys.* **15**, 884.
- Glatt, L. & Ellis, J. W. 1948 *J. Chem. Phys.* **16**, 551.
- Hutchinson, A. 1908 *Miner. Mag.* **15**, 93.
- Land, E. H. & West, C. D. 1946 *Colloid chemistry* (Alexander), **6**, 160. New York: Rheinhold Publishing Corporation.
- Mooney, R. C. L. 1941 *J. Amer. Chem. Soc.* **63**, 2828.
- Preston, J. M. 1933 *Modern textile microscopy*, p. 279. London: Emmott and Co.
- Sohon, F. W. 1941 *The stereographic projection*. New York: Chemical Publishing Co.
- Sutherland, G. B. B. M. & Vallance-Jones, A. 1947 *Nature*, **160**, 567.
- Thompson, H. W. & Mann, J. 1948 *Proc. Roy. Soc. A*, **192**, 489.
- Thompson, H. W. & Torkington, P. 1945 *Proc. Roy. Soc. A*, **184**, 21.
- Wright, N. 1948 *J. Opt. Soc. Amer.* **38**, 69.

Order-disorder statistics.

II. A two-dimensional model

By C. DOMB, *The Clarendon Laboratory, University of Oxford*

(Communicated by D. R. Hartree, F.R.S.—Received 1 February 1949)

The present paper is concerned with the detailed calculation of the partition functions for a two-dimensional quadratic lattice from the matrix derived previously. A brief discussion is given of the general methods of calculation available. The method here employed is the expansion of the eigenvector and eigenvalue as power series about zero temperature. This is readily applicable to finite matrices of quite high order, and, when a suitable notation has been introduced, to the infinite case. It is found more convenient to deal with the general problem of an unsymmetric net with different interactions in the two directions; for a ferromagnetic in the absence of a magnetic field a symmetry relation observed empirically enables terms to be derived successively, and it is assumed that this solution is equivalent to Onsager's (1944). The method is applied in the presence of a magnetic field, and several terms of a generalized series are deduced. Several terms are hence obtained of a series for the spontaneous magnetization. An inspection of the generalized series leads one to conjecture that the specific heat curve becomes continuous in the presence of a magnetic field. A rearrangement of the terms of the generalized series enables several terms of the high-temperature expansion to be deduced. Finally, the results are applied to the theory of binary solid solutions; the solubility curve for the two substances is formally closely related to the spontaneous magnetization. The separation into two phases is established, and the corresponding specific heat singularities are analyzed.

1. GENERAL DISCUSSION

In a previous paper (Domb 1949, hereafter referred to as I) it was shown that a number of problems in classical statistical mechanics could be reduced to the determination of the largest eigenvalues, λ , of infinite matrices of characteristic structure. When only nearest neighbour interactions are taken into account the elements of these matrices are functions of two variables, μ and z . [$z = \exp(-\epsilon_{12}/kT)$, ϵ_{12} being an interaction energy; $\mu = \exp(-2mH/kT)$ in the ferromagnetic problem, and is related to the concentration in the solid solution problem.] A brief discussion of the methods available for the evaluation of these eigenvalues follows.

(a) A well-known standard method is to raise the matrix X to a high power, and use the formula

$$\lambda = \lim_{p \rightarrow \infty} [\text{spur } X^p]^{1/p}. \quad (1)$$

The series expansions at low and high temperatures, described by Kramers & Wannier (1941), can be regarded as examples of the application of (1).

(b) The variation principle can be employed in a modified form applicable to non-Hermitian matrices.

(c) *Transformation.* In the absence of a magnetic field ($\mu = 1$) we have pointed out (I, §§ 3, 4) that the matrix X can be reduced to a V -matrix. For the two-dimensional nearest neighbour approximation, Kramers & Wannier discovered a transformation which enabled them to deduce that

$$\lambda(z) = \frac{1}{2}(1+z)^2 \lambda\left(\frac{1-z}{1+z}\right). \quad (2)$$

This was a very valuable piece of information, since it meant that the high-temperature development could be completely deduced from the low-temperature solution. It also enabled the singularity to be exactly located and given by

$$z = \frac{1-z}{1+z}. \quad (3)$$

We may note that the same transformation applied to the unsymmetric case of different intersections in different directions yields

$$\lambda(z, z') = \frac{1}{2}(1+z)(1+z')\lambda\left(\frac{1-z'}{1+z}, \frac{1-z}{1+z}\right). \quad (4)$$

This transformation has been considered topologically and extended by Onsager to apply to various two-dimensional nets (Wannier 1945), and it would seem worth investigating whether a generalization is possible for any of the other matrices described in I. Onsager (1944) suggests that the singular point in a case such as (4) is the point which transforms into itself, and is therefore given by

$$z = \frac{1-z'}{1+z'}, \quad z' = \frac{1-z}{1+z},$$

or

$$(1+z)(1+z') = 2. \quad (5)$$

(d) The eigenvector and eigenvalue can be expanded as power series about a certain point, and successive terms deduced by equating coefficients of like powers. This method is rather analogous to n th order perturbation theory of quantum mechanics. As in (a) we expand about points corresponding to zero temperature, and infinite temperature, since the zero-order matrices are very simple at these extremes. The method has been described and employed previously by Ashkin & Lamb* (1943). However, since it will be used extensively throughout the present paper, we shall discuss the process in some detail. It will be found that the method can readily be applied to finite matrices of quite high order, and by a suitable choice of notation to the limiting case when the order of the matrix becomes infinite. We now proceed to discuss the matter in greater detail.

2. EIGENVECTOR EXPANSION METHOD. FINITE MATRICES

We consider as an example the matrices corresponding to a two-dimensional strip in the absence of a magnetic field (I, § 3, $\mu = 1$). These matrices can be written in the form

$$A_m = A_0^{(m)} + A_1^{(m)}z + A_2^{(m)}z^2, \quad (6)$$

where $A_0^{(m)}, A_1^{(m)}, A_2^{(m)}$ are standard matrices of order 2^m with the following operational properties. (It will be convenient to use the scale of 2 for labelling components of

* I am indebted to Dr E. W. Montroll for drawing my attention to this interesting paper, which seems to have escaped notice in much of the literature.

a vector. Thus for a vector of order 2^m the general element will be denoted by $(a_1 \dots a_m)$, where each a_r is either 0 or 1.)

$$\left. \begin{aligned} A_0^{(m)}: & (0a_2 \dots a_{m-1} 0) \rightarrow (a_2 \dots a_{m-1} 00), \\ & (1a_2 \dots a_{m-1} 1) \rightarrow (a_2 \dots a_{m-1} 11), \\ A_1^{(m)}: & (0a_2 \dots a_{m-1} 1) \rightarrow (a_2 \dots a_{m-1} 10), \\ & (a_2 \dots a_{m-1} 11), \\ & (1a_2 \dots a_{m-1} 0) \rightarrow (a_2 \dots a_{m-1} 01), \\ & (a_2 \dots a_{m-1} 00), \\ A_2^{(m)}: & (0a_2 \dots a_{m-1} 0) \rightarrow (a_2 \dots a_{m-1} 01), \\ & (1a_2 \dots a_{m-1} 1) \rightarrow (a_2 \dots a_{m-1} 10). \end{aligned} \right\} \quad (7)$$

Each operator affects only components listed opposite it (for arbitrary $a_2 \dots a_{m-1}$), and gives zero for all other components.

In the absence of a magnetic field the operators are symmetric as regards interchange of 0 and 1. We shall confine our attention to the symmetric eigenvector, which is the case of practical interest; this is equivalent to halving the order of the matrices, since there are now 2^{m-1} independent components of the eigenvector. We shall usually list only one component, and the symmetric component is to be understood.

Let us now expand λ and $\mathbf{x}$ in the form

$$\lambda = \lambda_0 + \lambda_1 z + \lambda_2 z^2 + \dots + \lambda_p z^p + \dots, \quad (8)$$

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_1 z + \mathbf{x}_2 z^2 + \dots + \mathbf{x}_p z^p + \dots \quad (9)$$

Equating coefficients of successive powers of z in*

$$A\mathbf{x} = \lambda\mathbf{x}, \quad (10)$$

we obtain

$$\left. \begin{aligned} A_0 \mathbf{x}_0 &= \lambda_0 \mathbf{x}_0, \\ A_0 \mathbf{x}_1 + A_1 \mathbf{x}_0 &= \lambda_0 \mathbf{x}_1 + \lambda_1 \mathbf{x}_0, \\ A_0 \mathbf{x}_p + A_1 \mathbf{x}_{p-1} + A_2 \mathbf{x}_{p-2} &= \lambda_0 \mathbf{x}_p + \lambda_1 \mathbf{x}_{p-1} + \dots + \lambda_p \mathbf{x}_0. \end{aligned} \right\} \quad (11)$$

It is easy to show by raising A_0 to a high power that $\lambda_0 = 1$, and $\mathbf{x}_0$ has all components zero other than $(00 \dots 0)$, which we may conveniently take equal to 1. This corresponds to the completely ordered state. We note that $\mathbf{x}_0$ is also a left-hand eigenvector of A_0 , since

$$A'_0 \mathbf{x}_0 = \mathbf{x}_0. \quad (12)$$

Let us multiply the system (11) on the left-hand side by $\mathbf{x}_0$. Then in the first equation involving λ_p and $\mathbf{x}_p$, the terms in $\mathbf{x}_p$ cancel in virtue of (12). Hence if we know $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{p-1}$, we can determine λ_p . $\mathbf{x}_p$ is then given by

$$(1 - A_0)\mathbf{x}_p = (A_1 - \lambda_1)\mathbf{x}_{p-1} + (A_2 - \lambda_2)\mathbf{x}_{p-2} - \lambda_3 \mathbf{x}_{p-3} - \dots - \lambda_p \mathbf{x}_0 = \mathbf{u}_p. \quad (13)$$

Our determination of λ_p ensures the consistency of equation (13), and, in fact, the condition for this is that the $(00 \dots 0)$ component of $\mathbf{u}_p$ should be zero.

There is still some arbitrariness in the solution of (11), since any eigenvector can be multiplied by an arbitrary factor. We shall find it convenient to remove this

* For convenience the superscript m is dropped, and it is to be understood in all that follows.

arbitrariness by taking the $(00 \dots 0)$ component of $\mathbf{x}$ to be unity for all z . This means that the $(00 \dots 0)$ component of $\mathbf{x}_p$ is zero for $p > 0$, and

$$\mathbf{x}'_0 \mathbf{x}_p = 0 \quad (p > 0). \quad (14)$$

Multiplying (11) on the left-hand side by $\mathbf{x}_0$ we find that λ_p is the $(100 \dots 0)$ component of $\mathbf{x}_{p-1}$. The vector $\mathbf{x}_p$ is then determined from (13) as

$$\mathbf{x}_p = (1 + A_0 + A_0^2 + \dots + A_0^{m-2}) \mathbf{u}_p, \quad (15)$$

since $A_0^n \mathbf{u}_p$ is zero for $n \geq m-1$.

It is now a straightforward task to evaluate several terms of the expansion for finite matrices even of quite high order, and an example is given in table 1* of the application of the method to the case $m = 5$. We find

$$\lambda^{(5)} = 1 + z^4 + 3z^6 + 15z^8 + \dots \quad (16)$$

The process becomes more laborious as the order of the terms increases, but there is a well-known short cut which acts as a useful check on the numerical correctness of the eigenvalue, and enables us to double the number of terms of the eigenvalue.

Let us try to determine the left-hand eigenvector, $\mathbf{y}$, of A which satisfies the equation

$$A' \mathbf{y} = (A'_0 + A'_1 z + A'_2 z^2) \mathbf{y} = \lambda \mathbf{y}. \quad (17)$$

The procedure is analogous to that described above, except that the basic operators now have the following properties:

$$\left. \begin{aligned} A'_0: & \begin{aligned} (a_2 \dots a_{n-1} \ 00) & \rightarrow (0a_2 \dots a_{n-1} \ 0), \\ (a_2 \dots a_{n-1} \ 11) & \rightarrow (1a_2 \dots a_{n-1} \ 1), \end{aligned} \\ A'_1: & \begin{aligned} (a_2 \dots a_{n-1} \ 10) + (a_2 \dots a_{n-1} \ 11) & \rightarrow (0a_2 \dots a_{n-1} \ 1), \\ (a_2 \dots a_{n-1} \ 01) + (a_2 \dots a_{n-1} \ 00) & \rightarrow (1a_2 \dots a_{n-1} \ 0), \end{aligned} \\ A'_2: & \begin{aligned} (a_2 \dots a_{n-1} \ 01) & \rightarrow (0a_2 \dots a_{n-1} \ 0), \\ (a_2 \dots a_{n-1} \ 10) & \rightarrow (1a_2 \dots a_{n-1} \ 1). \end{aligned} \end{aligned} \right\} \quad (18)$$

λ_p is now given as the $(00 \dots 01)$ component of $\mathbf{y}_{p-2}$, and $\mathbf{y}_p$ by a relation similar to (15). Both methods should lead to the same value of λ_p , and this can be used as a check.

Suppose we have now evaluated p terms of one eigenvector and q terms of the other, and let

$$\left. \begin{aligned} \xi_p &= \mathbf{x}_0 + \mathbf{x}_1 z + \dots + \mathbf{x}_p z^p = \mathbf{x} + O(z^{p+1}), \\ \eta_q &= \mathbf{y}_0 + \mathbf{y}_1 z + \dots + \mathbf{y}_q z^q = \mathbf{y} + O(z^{q+1}). \end{aligned} \right\} \quad (19)$$

Then it is easy to show that (Aitken 1937)

$$\nu_{pq} = \frac{\eta'_q A \xi_p}{\eta'_q \xi_p} = \lambda + O(z^{p+q+2}), \quad (20)$$

thus giving $(p+q+1)$ terms of the eigenvalue correctly.

Applying this to the case $m = 5$ we deduce the additional terms

$$+ 48z^{10} + 70z^{12} - 253z^{14}. \quad (21)$$

* The tables are printed at the end of the paper.

The high-temperature expansion can be discussed similarly in powers of $1-z$, or $(1-z)/(1+z)$, with different operators replacing A_0 , A_1 and A_2 . However, the transformation (3) enables us to avoid this discussion, and write down the appropriate expansion immediately.

3. EIGENVECTOR EXPANSION METHOD. INFINITE MATRICES

The method of the last section is only practicable when all components of the eigenvectors required can be written down in order. When we come to consider large values of m this becomes impossible. The method can still be applied, but a suitable notation to cover classes of components with similar properties must first be chosen.

The notation to be employed can perhaps best be understood by reference to a few examples. A suffix after 0 or 1 denotes the number of times the 0 or 1 is repeated; the sum of all the suffixes in any configuration must be equal to m . The configuration $0_r 1_0$ will be used to denote the $(m-2)$ components from $r=1$, $s=m-2$ to $r=m-2$, $s=1$. $0_r 1_0 1_0$ denotes all components such that

$$r \geq 1, \quad s \geq 1, \quad t \geq 1, \quad r+s+t = m-2;$$

there are $\frac{1}{2}(m-3)(m-4)$ such components. Similarly $r, s, t, \dots$ will in general be used as variable suffixes, whereas $a, b, c, \dots$ will be fixed suffixes. $1_a 0_r 1_b 0_s 1_c 0_t$ denotes all components such that

$$r \geq 1, \quad s \geq 1, \quad t \geq 1, \quad r+s+t = m - (a+b+c).$$

As typical configurations we may write down

$$0_r 1_3, \quad 1_2 0_r 1_0, \quad 0_r 1_2 0_s 1_0, \quad 1_0 1_2 0_s 1.$$

It will be noted that 1's always occur with fixed suffixes. But we shall sometimes also require 0's with fixed suffixes, and the same convention will be used. Thus $0_r 1_0 0$ denotes all components such that

$$r \geq 1, \quad s \geq 1, \quad r+s = m-2;$$

this consists of all components of $0_r 1_0$ other than $0_r 1_0$, and could alternatively be written $0_r 1_0 - 0_r 1_0$. With a little experience one rapidly becomes quite accustomed to the notation.

The zero-order vector $\mathbf{x}_0$ contains only one non-zero component $0_r = 1$, and for all other vectors $\mathbf{x}_p$ this component is zero. A list will be given of the non-zero configurations of the $\mathbf{x}_p$, and if no number is put opposite a configuration unity is to be understood. It has been shown in the last section that λ_p is equal to the 1_0 component of $\mathbf{x}_{p-1}$. The procedure for determining $\mathbf{x}_p$ from $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{p-1}$ consists in writing down $\mathbf{u}_p$ from (13), and then operating on $\mathbf{u}_p$ with $g(A_0)$, by (15), where

$$g(A_0) = 1 + A_0 + A_0^2 + \dots + A_0^{m-2}. \quad (22)$$

Table 2 shows the first eight vectors which have been determined by this process, and enables one to write down λ_1 to λ_8 . The procedure is quite straightforward once

one has acquired facility in applying the operators A_0 , A_1 and A_2 to the various configurations, using (8). The following are some examples of these operations:

$$\left. \begin{aligned} A_2: \quad (0_r 10_s) &\rightarrow \left. \begin{aligned} 0_r 10_s 1 \\ 10_r 1 \end{aligned} \right\}, \\ g(A_0): \quad (0_r 10) &\rightarrow \left. \begin{aligned} 0_r 10_s \\ 10_r \end{aligned} \right\}, \\ A_1: \quad (0_r 10_s 1) &\rightarrow \left. \begin{aligned} 0_r 10_s 10 & 0_r 10_s 1_2 \\ 10_r 10 & 10_r 1_2 \end{aligned} \right\}, \\ A_1: \quad (10_r 10_s) &\rightarrow \left. \begin{aligned} 0_r 10_s 0 \\ 0_r 10_s 1 \end{aligned} \right\}, \\ A_2: \quad (10_r 1) &\rightarrow 0_r 10 \\ g(A_0): \quad (10_r 1) &\rightarrow \left. \begin{aligned} 10_r 1 \\ 0_r 1_2 \end{aligned} \right\}. \end{aligned} \right\} \quad (23)$$

The derivation of table 2 may be illustrated by giving the details of the determination of $\mathbf{x}_5$ from $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ and $\mathbf{x}_4$.

$$\begin{array}{rclcl} A_1 \mathbf{x}_4 & 0_r 10 & 0_r 1_2 & 0_r 1_3 0 & 0_r 1_4 \\ & 10_r & 10_r 1 & & \\ A_2 \mathbf{x}_3 & & 0_r 10_s 1 & & \\ & & 10_r 1 & & \\ \mathbf{u}_5 = A_1 \mathbf{x}_4 + A_2 \mathbf{x}_3 - \mathbf{x}_1 & & & & \end{array}$$

$\mathbf{x}_5$ is then given in table 2 by $g(A_0) \mathbf{u}_5$.

The terms of the left-hand eigenvector can be determined in an analogous manner using (18), and table 3 gives the first four vectors. New configurations appear very much more rapidly than with the right-hand eigenvector, and the complication is correspondingly increased. At any stage we can apply (20) to obtain more terms of the eigenvalue, but the calculation is rather tedious.

Some empirical observations may be made on the properties of the terms of the eigenvectors. Any given component occurs either only in odd terms or only in even terms; hence in the power series corresponding to this component successive terms differ by z^2 . The eigenvectors of the matrix T_m (I, equation (18)) are symmetric about the centre, so that the $(a_1 \dots a_m)$ component is always equal to the $(a_m \dots a_1)$ component. Also T_m can easily be symmetrized (Kramers & Wannier 1941), and there is thus a simple relation between its left-hand and right-hand eigenvectors; if ν is the number of unlike links (0 to 1 or 1 to 0) in $(a_1 \dots a_m)$, then the ratio of the $(a_1 \dots a_m)$ component of the right-hand eigenvector to the $(a_1 \dots a_m)$ component of the left-hand eigenvector is z^ν . Even though the screw transformation destroys the central symmetry of the eigenvector, a curious relationship persists between the left- and right-hand eigenvectors which enables one to be written down if the other is known. It will be found on inspection that the ratio of the $(a_1 \dots a_m)$ component of the right-hand eigenvector to the $(a_m \dots a_1)$ component of the left-hand eigenvector is z^ν . This latter result applies only to the infinite case, and not to any finite matrices.

It will be seen that the configurations in $\mathbf{x}_p$ and $\mathbf{y}_p$ do not enter in any simple systematic manner, and high-order terms become extremely laborious to evaluate. It will be found that a considerable simplification can be effected by dealing with the more general problem of a net with unequal interactions in the two directions, and we shall now discuss this in further detail.

4. THE UNSYMMETRIC NET

We have already considered the formulation of the matrix appropriate to the unsymmetric net (I, § 3). Instead of (7) we can now write

$$A = A_{00} + A_{01}z + A_{10}z' + A_{11}z'z, \quad (24)$$

where A_{00} is identical with A_0 , A_{11} with A_2 , but A_{01} and A_{10} have the following operational properties:

$$\left. \begin{aligned} A_{01}: \quad & (0a_2 \dots a_{m-1} 1) \rightarrow (a_2 \dots a_{m-1} 10), \\ & (1a_2 \dots a_{m-1} 0) \rightarrow (a_2 \dots a_{m-1} 01), \\ A_{10}: \quad & (0a_2 \dots a_{m-1} 1) \rightarrow (a_2 \dots a_{m-1} 11), \\ & (1a_2 \dots a_{m-1} 0) \rightarrow (a_2 \dots a_{m-1} 00). \end{aligned} \right\} \quad (25)$$

It will be seen that whereas (7) is quadratic in z , (24) is linear in both z and z' .

The operators arising in the present work can be divided into two classes, 'shift' operators such as A_{00} and A_{01} , and 'creation and annihilation' operators such as A_{10} and A_{11} . Part of the complication of the preceding section arose because the operator A_1 combines both types of operation, whereas in this section each operator belongs to only one class. We shall see later that the presence of a magnetic field further differentiates between the 'creation' and 'annihilation' operations, and between the types of 'shift' operation.

We must now expand λ and $\mathbf{x}$ as double power series in z and z'

$$\left. \begin{aligned} \lambda &= \sum_{pq=0}^{\infty} \lambda_{pq} z'^p z^q, \\ \mathbf{x} &= \sum_{pq=0}^{\infty} \mathbf{x}_{pq} z'^p z^q. \end{aligned} \right\} \quad (26)$$

We know on physical grounds that λ must be symmetric in z and z' , so that

$$\lambda_{pq} = \lambda_{qp}. \quad (27)$$

This is of course a property only of the infinite strip. Also the vector $\mathbf{x}$ will be found to be quite unsymmetrical in z and z' .

We can now substitute from (26) into (10), equate coefficients of the corresponding powers, and proceed as previously. The result for the first few terms is shown in table 4. We can, on inspection, find considerable regularity about the occurrence of various configurations, and for the first few rows we can even write down the general term. The vectors $\mathbf{x}_p$ of the preceding section correspond to the case $z' = z$, and can be obtained by diagonal summation,

$$\mathbf{x}_p = \sum_{t=0}^p \mathbf{x}_{t(p-t)}. \quad (28)$$

One thus sees the reason for the complication in the types and values of configurations constituting $\mathbf{x}_p$.

Further progress can be made by summing (26) with respect to z' or z . We then have

$$\left. \begin{aligned} \lambda &= \sum_{p=0}^{\infty} \mu_p z'^p, \\ \mathbf{x} &= \sum_{p=0}^{\infty} \zeta_p z'^p, \end{aligned} \right\} \quad (29)$$

μ_p and ζ_p now being functions of z . We may write (24) in the form

$$\text{where} \quad \left. \begin{aligned} A &= T_0 + T_1 z', \\ T_0 &= A_{00} + A_{01} z, \\ T_1 &= A_{10} + A_{11} z. \end{aligned} \right\} \quad (30)$$

We have summed with respect to the variable which causes the two shift operators to be paired together in T_0 ; experience shows that this is the more useful choice.

The properties of the operators T_0 and T_1 can readily be established with the aid of relations similar to (23). For $z < 1$, $T_0^n \tau$ tends to zero as $n \rightarrow \infty$ for any vector τ whose 0_r component is zero. Hence by analogy with (13) and (15) we deduce that

$$\left. \begin{aligned} \zeta_p &= f(T_0) \tau_p, \\ \tau_p &= (T_1 - \mu_1) \zeta_{p-1} - \mu_2 \zeta_{p-2} - \dots - \mu_p \zeta_0 \\ \text{where} \quad f(T_0) &= \sum_{q=0}^{\infty} T_0^q. \end{aligned} \right\} \quad (31)$$

The operation of $f(T_0)$ is somewhat different from anything previously encountered, and some examples of its application follow:

$$\left. \begin{aligned} f(T_0)(0_r 1) &\rightarrow \left\{ \begin{aligned} 0_r 1 &= 1 + z^2 + z^4 + z^6 + \dots = 1/(1 - z^2), \\ 0_r 10_s &= z + z^3 + z^5 + \dots = z/(1 - z^2), \\ 10_r &= z + z^3 + \dots = z/(1 - z^2), \end{aligned} \right. \\ f(T_0)(1_2 0_r) &\rightarrow \left\{ \begin{aligned} 1_2 0_r &= 1/(1 - z^2), \\ 10_r 1 &= z/(1 - z^2), \\ 0_r 1_2 &= z/(1 - z^2), \\ 0_r 1_2 0_s &= z^2/(1 - z^2), \end{aligned} \right. \\ f(T_0)(10_r 1_2 0_s) &\rightarrow \left\{ \begin{aligned} 10_r 1_2 0_s &= 1/(1 - z^4), \\ 0_r 1_2 0_s 1 &= z/(1 - z^4), \\ 0_r 1_2 0_s 10_t &= z^2/(1 - z^4), \\ 1_2 0_r 10_s &= z^2/(1 - z^4), \\ 10_r 10_s 1 &= z^3/(1 - z^4), \\ 0_r 10_s 1_2 &= z^3/(1 - z^4), \\ 0_r 10_s 1_2 0_t &= z^4/(1 - z^4). \end{aligned} \right. \end{aligned} \right\} \quad (32)$$

Table 5 shows the first few vectors ζ_p and the corresponding terms μ_p . The configurations of each vector can be collected in groups with equal numbers of 1's, and each vector contains either odd or even numbers; thus ζ_3 contains only configurations

with one or three 1's, ζ_4 with two or four, ζ_5 with one, three or five, and so on. If we are interested only in the terms of the eigenvalue, μ_p , it is only necessary to calculate a portion of the eigenvector from a certain stage onwards, and this introduces a considerable simplification. Thus for μ_6 we need only the two 1's group of ζ_4 and the single one group of ζ_5 ; for μ_8 the one and three group of ζ_5 , the two group of ζ_6 , and the single group of ζ_7 . This is one of the advantages of combining the two shift operators to give T_0 in (30). It corresponds to the fact that we can immediately write down the left-hand eigenvector χ of A to the same order of magnitude as we have determined ζ ; in fact, with the same notation as at the end of § 3, the $(a_1 \dots a_m)$ component of ζ_p is equal to z^p times the $(a_m \dots a_1)$ component of χ_p .

It will be seen from (30) that configurations are formed in particular groups with factors such as $1/(1-z^2)$, $1/(1-z^4)$, $1/(1-z^6)$, ... We note empirically, however, that the factors other than $1/(1-z^2)$ cancel out, and that any configuration of ζ_p can be expressed as $1/(1-z^2)^p$ multiplied by a polynomial.

There is also a curious symmetry about the polynomials corresponding to the components $(a_1 \dots a_m)$ and $(a_m \dots a_1)$ which will be recognized on writing down a few examples.

$$\begin{aligned}
 (1-z^2)^3 \zeta_3: \quad & 1_2 0_r 1 = 2z^3 + 3z^5, & 10_r 1_2 &= 3z^3 + 2z^5, \\
 & 1_3 0_r = z^2 + 3z^4 + z^6, & 0_r 1_3 &= z + 3z^3 + z^5, \\
 (1-z^2)^4 \zeta_4: \quad & 1_2 0_r = z^2 + 3z^4, & 0_r 1_2 &= 3z^5 + z^7, \\
 & 10_r 10_s 1_2 = 3z^5 + 2z^7, & 1_2 0_r 10_s 1 &= 2z^5 + 3z^7, \\
 (1-z^2)^5 \zeta_5: \quad & 10_r = z^2 + z^4 + 2z^6, & 0_r 1 &= 2z^5 + z^7 + z^9, \\
 & 1_2 0_r 1 = z^2 + 7z^5 + 5z^7 + z^9, & 10_r 1_2 &= z^3 + 5z^5 + 7z^7 + z^9. \quad (33)
 \end{aligned}$$

In general terms if the two components are given by

$$\begin{aligned}
 (a_1 \dots a_m) &= h(z, z'), \\
 (a_m \dots a_1) &= k(z, z'),
 \end{aligned} \quad (34)$$

then

$$h(z, z') = \pm z^p k\left(\frac{1}{z}, z'\right), \quad (35)$$

where the positive or negative sign is to be taken according as the number of 1's in the configuration is even or odd. It has not been found possible to establish the relation (35) rigorously; but there seems little doubt that it is correct.

We shall find that the application of (35) to the $0_r 1$ and 10_r configurations enables us to derive a general method of obtaining successive terms μ_p . Let us write

$$\begin{aligned}
 \mu_{p+1} &= [10_r]_p = l_p(z)/(1-z^2)^p, \\
 [0_r 1]_p &= m_p(z)/(1-z^2)^p.
 \end{aligned} \quad (36)$$

From table 5 we then have

$$\left. \begin{aligned}
 l_1(z) &= z^2, & m_1(z) &= z, \\
 l_3(z) &= z^2, & m_3(z) &= z^5, \\
 l_5(z) &= z^2 + z^4 + 2z^6, & m_5(z) &= 2z^5 + z^7 + z^9,
 \end{aligned} \right\} \quad (37)$$

and similarly

$$l_7(z) = z^2 + 3z^4 + 11z^6 + 7z^8 + 3z^{10}, \quad m_7(z) = 3z^5 + 7z^7 + 11z^9 + 3z^{11} + z^{13}.$$

It is assumed that each successive term in the series has two additional terms in the corresponding polynomial. The symmetry property (27) of λ enables one to write down the first k terms of the polynomial $l_{(2k+1)}$ when the previous polynomials are all known. The second equation of the set (10) gives a relation between (10_r) and $(0_r, 1)$ which enables the remaining $(k-1)$ terms to be deduced. This equation is

$$zz'0_r + z(10_r) = \lambda(0_r, 1), \quad (38)$$

and substituting the power series for λ , $0_r, 1$ and 10_r and comparing coefficients of like powers of z' , we find that

$$\left. \begin{aligned} m_1(z) &= z l_1(z) + z(1 - z^2), \\ m_3(z) &= z l_3(z) - (1 - z^2) l_1 m_1, \\ m_5(z) &= z l_5(z) - (1 - z^2) (l_1 m_3 + l_3 m_1), \\ m_7(z) &= z l_7(z) - (1 - z^2) (l_1 m_5 + l_3 m_3 + l_5 m_1), \end{aligned} \right\} \quad (39)$$

and so on. It will be found that the polynomials (37) satisfy the relations (39), and we can quite easily deduce more terms of the series:

$$\left. \begin{aligned} l_9(z) &= z^2 + 6z^4 + 36z^6 + 63z^8 + 64z^{10} + 22z^{12} + 4z^{14}, \\ l_{11}(z) &= z^2 + 10z^4 + 90z^6 + 297z^8 + 562z^{10} + 504z^{12} + 245z^{14} + 50z^{16} + 5z^{18}, \\ l_{13}(z) &= z^2 + 15z^4 + 190z^6 + 1002z^8 + 3061z^{10} + 5038z^{12} + 4805z^{14} + 2496z^{16} \\ &\quad + 715z^{18} + 95z^{20} + 6z^{22}. \end{aligned} \right\} \quad (40)$$

It is of some interest to tabulate the first few terms of the double-power series for λ , and these are given in table 6.

We presume that power-series solution derived here is identical with that given in closed form in terms of elliptic integrals by Onsager (1944). The power series agree as far as all the terms given by Onsager, although the complete identification would appear to present some mathematical difficulty. The high-temperature expansion is given by the transformation (4), and hence the complete solution is available for this case. The physical details of the solution are not considered here, since they have already been discussed at length by Onsager.

5. PRESENCE OF A MAGNETIC FIELD

When a magnetic field is present the eigenvalue is a function of two variables μ and z , and all practical cases are covered by considering values of these variables between 0 and 1. A capital letter Λ will now be used to denote the eigenvalue. The case $\mu > 1$ is taken account of by the relation (I, end of § 2)

$$\mu^{-1}\Lambda(\mu) = \mu^{\frac{1}{2}}\Lambda(1/\mu), \quad (41)$$

and the case $z > 1$ corresponds to an antiferromagnetic which is not discussed here. Before proceeding with the detailed mathematics it is useful to give a brief physical account of the salient features and requirements in the presence of a magnetic field.

The region of interest is represented by the square $ABED$ in figure 1. Onsager's solution in the absence of a magnetic field corresponds to the path AB , and C is the Curie point $z = 0.414$. Our first requirement is the location of the singularities of the eigenvalue Λ in μ and z . Along the line AC we expect a singularity in Λ as a function of μ , giving rise to the spontaneous magnetization as $\mu \rightarrow 1$ (that is, as the magnetic field tends to zero). At C we have a singularity in λ as a function of z (the specific heat singularity). We should not expect any singularity along the portion CB of the curve, since the spontaneous magnetization has then disappeared. The specific heat curve of a ferromagnetic in the presence of a constant non-zero magnetic field is given by a path from D to B of the form

$$\mu = z^k. \quad (42)$$

Small k correspond to weak magnetic fields, and large k to strong fields.

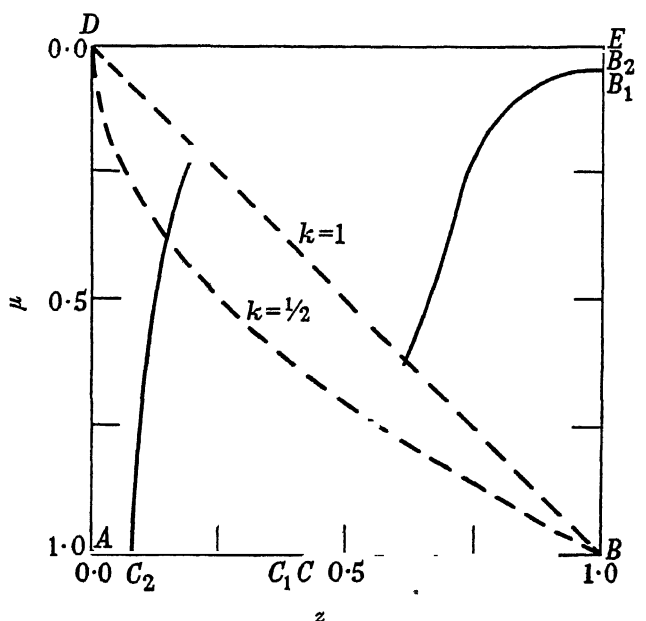


FIGURE 1

We shall attempt to find a power-series solution, as before, but our series is now a double-power series in μ and z , and there are alternative ways of looking at it. For example, we could start at any point of AD and proceed along a path parallel to AB obtaining a power series in z whose coefficients are functions of μ . Alternatively, we could start at some point of DE and proceed along a path parallel to DA obtaining a power series in μ whose coefficients are functions of z . Figure 1 gives a useful indication of where it is feasible to start power-series expansions. For example, in

considering finite matrix approximations, it would be inadvisable to start at any point of AC because of the singularity as a function of μ , but any point of CB should yield useful information more readily.

We shall start with the first alternative, but will consider, as in the last section, the general problem of a net with different interactions in the two directions which gives rise to a double-power series, symmetric in z and z' , and with coefficients functions of μ . By analogy with (29) and (26)

$$\left. \begin{aligned} \Lambda &= \sum_{p=0}^{\infty} M_p z'^p = \sum_{pq=0}^{\infty} \Lambda_{pq} z'^p z^q, \\ X &= \sum_{p=0}^{\infty} Z_p z'^p, \end{aligned} \right\} \quad (43)$$

where capital letters have been used to indicate that the quantities concerned are functions of μ . Any operator encountered previously is now a function of μ , and can be written in the form

$$T = T^+ + \mu T^-. \quad (44)$$

The basic operators ((24) and (25)) have the following properties:

$$\left. \begin{aligned} A_{00}^+ &: (0a_2 \dots a_{m-1} 0) \rightarrow (a_2 \dots a_{m-1} 00), & A_{00}^- &: (1a_2 \dots a_{m-1} 1) \rightarrow (a_2 \dots a_{m-1} 11), \\ A_{01}^+ &: (0a_2 \dots a_{m-1} 1) \rightarrow (a_2 \dots a_{m-1} 10), & A_{01}^- &: (1a_2 \dots a_{m-1} 0) \rightarrow (a_2 \dots a_{m-1} 01), \\ A_{10}^+ &: (1a_2 \dots a_{m-1} 0) \rightarrow (a_2 \dots a_{m-1} 00), & A_{10}^- &: (0a_2 \dots a_{m-1} 1) \rightarrow (a_2 \dots a_{m-1} 11), \\ A_{11}^+ &: (1a_2 \dots a_{m-1} 1) \rightarrow (a_2 \dots a_{m-1} 10), & A_{11}^- &: (0a_2 \dots a_{m-1} 0) \rightarrow (a_2 \dots a_{m-1} 01). \end{aligned} \right\} \quad (45)$$

The procedure is very similar to that of the last section, and by analogy with (32) the following are some examples of the operation of $f(T)$:

$$\left. \begin{aligned} f(T_0)(0_r 1) &\rightarrow \left\{ \begin{aligned} 0_r 1 &= 1 + \mu z^2 + \mu^2 z^4 + \dots = 1/(1 - \mu z^2), \\ 0_r 10_s &= z + \mu z^3 + \mu^2 z^5 + \dots = z/(1 - \mu z^2), \\ 10_r &= z + \mu z^3 + \mu^2 z^5 + \dots = z/(1 - \mu z^2), \end{aligned} \right. \\ f(T_0)(1_2 0_r) &\rightarrow \left\{ \begin{aligned} 1_2 0_r &= 1/(1 - \mu^2 z^2), \\ 10_r 1 &= \mu z/(1 - \mu^2 z^2), \\ 0_r 1_2 &= \mu^2 z/(1 - \mu^2 z^2), \\ 0_r 1_2 0_s &= \mu^3 z^2/(1 - \mu^2 z^2), \end{aligned} \right. \\ f(T_0)(10_r 1_2 0_s) &\rightarrow \left\{ \begin{aligned} 10_r 1_2 0_s &= 1/(1 - \mu^3 z^4), \\ 0_r 1_2 0_s 1 &= \mu z/(1 - \mu^3 z^4), \\ 0_r 1_2 0_s 10_t &= \mu z^2/(1 - \mu^3 z^4), \\ 1_2 0_r 10_s &= \mu z^2/(1 - \mu^3 z^4), \\ 10_r 10_s 1 &= \mu^2 z^3/(1 - \mu^3 z^4), \\ 0_r 10_s 1_2 &= \mu^3 z^3/(1 - \mu^3 z^4), \\ 0_r 10_s 1_2 0_t &= \mu^3 z^4/(1 - \mu^3 z^4). \end{aligned} \right. \end{aligned} \right\} \quad (46)$$

Table 7 shows the first few terms of the eigenvector, and from this we deduce the first three non-zero coefficient in the eigenvalue:

$$\begin{aligned}
 M_2 &= \mu z^2 / W_1, \\
 M_4 &= \frac{\mu^2 z^2 + (\mu^2 + 2\mu^3) z^4}{W_1^2 W_2} - \frac{3\mu^2}{W_1^3}, \quad W_r = 1/(1 - \mu^r z^2), \\
 M_6 &= \frac{\mu^3 z^2 + (3\mu^4 + 2\mu^5) z^4 + (3\mu^5 + 4\mu^6) z^6 + (\mu^6 + 2\mu^7) z^8}{W_1^2 W_2^2 W_3} \\
 &\quad + \frac{2(\mu^3 + \mu^4) z^4 + 2(\mu^4 + \mu^5) z^6}{W_1^2 W_2 W_3} + \frac{\mu^5 z^6}{W_1^4 W_3} \\
 &\quad + \frac{(-3\mu^3 + \mu^4) z^4 + (-4\mu^4 + 4\mu^5) z^6 + (-\mu^5 + 3\mu^6) z^8}{W_1^3 W_2^2} \\
 &\quad - \frac{6\mu^3 z^4 + (8\mu^3 + 18\mu^4) z^6 + (2\mu^4 + 6\mu^5) z^8}{W_1^4 W_2} + \frac{16\mu^3 z^6 + 3\mu^4 z^8}{W_1^5}.
 \end{aligned} \tag{47}$$

When we compare (47) with (36) and (37) we see the enormous increase in complication which the magnetic field produces. If we have a series of coefficients such as (47) we can obtain the best approximation by using the symmetry properties of Λ in z and z' . Thus if

$$\left. \begin{aligned}
 \Lambda^{(r)}(z, z') &= \sum_{p=0}^r M_p z'^p \\
 \Lambda^{(r,r)}(z, z') &= \sum_{p,q=0}^r \Lambda_{pq} z'^p z^q,
 \end{aligned} \right\} \tag{48}$$

and

the sum of all terms in the first r rows and columns is

$$\Lambda^{(r)}(z, z') + \Lambda^{(r)}(z', z) - \Lambda^{(r,r)}(z, z'). \tag{49}$$

It is interesting to write down the first terms of the power series in z corresponding to the symmetric case $z = z'$, and these can easily be deduced from (47). We obtain

$$\begin{aligned}
 \Lambda &= 1 + \mu z^4 + 2\mu^2 z^6 + (-2\mu^2 + 6\mu^3 + \mu^4) z^8 + (-14\mu^3 + 18\mu^4 + 8\mu^5 + 2\mu^6) z^{10} \\
 &\quad + (8\mu^3 - 77\mu^4 + 44\mu^5 + 40\mu^6 + 22\mu^7 + 6\mu^8 + \mu^9) z^{12} \\
 &\quad + (98\mu^4 - 370\mu^5 + 40\mu^6 + 138\mu^7 + 134\mu^8 + 72\mu^9 + 30\mu^{10} + 8\mu^{11} + 2\mu^{12}) z^{14} + \dots
 \end{aligned} \tag{50}$$

An important feature of the series deduced in this section is that they seem to be valid in μ up to $\mu = 1$; they can therefore be used to determine the value of Λ_μ as $\mu \rightarrow 1$, and hence the spontaneous magnetization. The series (50) differs essentially from those corresponding to finite matrices in that the coefficients of the powers of z are finite polynomials in μ and not infinite series. For any finite strip at some stage the coefficients would become infinite series, and this corresponds to the physical fact that a finite number of overturned spins could completely separate the strip into two parts. (This can be verified, for example, in the one-dimensional solution, I, equation (15).)

We shall have occasion to manipulate the series (50) considerably in the next two sections. Since only a finite number of terms have been derived, it is impossible to justify these processes rigorously. It is believed, however, that one can generally judge on physical grounds whether the results obtained are valid.

6. THE SPONTANEOUS MAGNETIZATION*

If we differentiate (50) with respect to μ and let $\mu \rightarrow 1^-$, we find that $\Lambda_\mu/\Lambda \rightarrow \beta(z)$, where

$$\lambda(z)\beta(z) = z^4 + 4z^6 + 18z^8 + 82z^{10} + 387z^{12} + 1880z^{14} + \dots \quad (51)$$

The relation (41) enables us to obtain the values of Λ for $\mu > 1$ from those for $\mu < 1$, and it is easy to show that as $\mu \rightarrow 1^+$, $\Lambda_\mu/\Lambda \rightarrow 1 - \beta(z)$. Hence, since $\beta(z) \neq \frac{1}{2}$, there is a discontinuity in Λ_μ at $\mu = 1$. (This indicates that the series (50) can only be valid for $\mu \leq 1$.)

When we take account of the zero energy terms, the complete partition function for a ferromagnetic is given by

$$Z = \mu^{-\frac{1}{2}} z^{-1} \Lambda(\mu, z), \quad (52)$$

and the free energy, F , is equal to $-kT \log Z$. The magnetization is given by $-\partial F/\partial H$, and hence by

$$I/I_\infty = 1 - 2\mu\Lambda_\mu/\Lambda. \quad (53)$$

Thus the discontinuity in Λ_μ gives rise to a spontaneous magnetization

$$I_0/I_\infty = 1 - 2\beta(z). \quad (54)$$

As before we can more easily deal with the general problem of an unsymmetric net, whose spontaneous magnetization is a function of z and z' . Suppose that $\Lambda_\mu \rightarrow \alpha(z, z')$ as $\mu \rightarrow 1$, and

$$\alpha(z, z') = \sum_{p=0}^{\infty} z'^p s_p(z)/(1-z^2)^p = \sum_{pq=0}^{\infty} \alpha_{pq} z'^p z^q. \quad (55)$$

From (47) by differentiating with respect to μ we deduce that

$$\left. \begin{aligned} s_2(z) &= z^2, \\ s_4(z) &= 2z^2 + 4z^4 + z^6, \\ s_6(z) &= 3z^2 + 19z^4 + 32z^6 + 15z^8 + z^{10}. \end{aligned} \right\} \quad (56)$$

The terms (51) and (56) can be deduced from the complete eigenvector solution in the absence of a magnetic field. This can be more readily demonstrated by dealing with the terms of $\beta(z)$ and $\beta(z, z')$ directly. Writing

$$\beta(z, z') = \sum_{p=0}^{\infty} z'^p t_p(z)/(1-z^2)^p = \sum_{pq=0}^{\infty} \beta_{pq} z'^p z^q, \quad (57)$$

we readily find that

$$\left. \begin{aligned} t_2(z) &= z^2, \\ t_4(z) &= 2z^2 + 3z^4 + 2z^6, \\ t_6(z) &= 3z^2 + 16z^4 + 32z^6 + 16z^8 + 3z^{10}, \end{aligned} \right\} \quad (58)$$

$$\text{and} \quad \beta(z) = z^4 + 4z^6 + 17z^8 + 76z^{10} + 357z^{12} + 1736z^{14} + \dots \quad (59)$$

It will be observed that the $t(z)$ polynomials are symmetric about the central term.

If the series $1 - 2\beta(z)$ is now compared with the series for long range order in Ashkin & Lamb's paper (1943, equation 5.10) it will be found that they are identical for all terms quoted. It is not difficult to see that the spontaneous magnetization and long range order must be identical since they are both given by $\overline{n_1 - n_2}$. The methods used by Ashkin & Lamb for deducing the series for long range order can be

* I am indebted to a referee for criticism of the original draft of this section.

applied to table 5, and the three terms of (58) result. The next term of this series has also been deduced similarly:

$$t_4(z) = 4z^2 + 46z^4 + 200z^6 + 305z^8 + 200z^{10} + 46z^{12} + 4z^{14}, \quad (60)$$

and this adds to (59) the two terms

$$8659z^{16} + 44024z^{18}. \quad (61)$$

One of the interesting physical properties of the series $\beta(z)$ is its behaviour in the neighbourhood of the singularity, and this involves an estimation of the magnitude and trend of the terms (59), (61). It is useful as a preliminary to quote briefly the properties of a power series $\sum a_n z^n$ whose coefficients a_n are steady in behaviour (Dienes 1931). If the a_n are consistent in sign there is a singularity on the real axis at $z = R$, the radius of convergence of the series. $1/R$ is equal to the limit as $n \rightarrow \infty$ of $|a_n|^{1/n}$, and the behaviour of the function in the neighbourhood of the singularity is determined by the order of magnitude of $R^n a_n$. For example, if $R^n a_n \sim K$ the singularity is of the form $(1 - z/R)^{-1}$, if $R^n a_n \sim 1/n$ it is of the form $\log(1 - z/R)$. If $R^n a_n \sim A/n^k$ an estimate of the value of k for large n enables one to obtain a good approximation to the sum of the series. For the series (59) one can feel fairly confident on physical grounds that the radius of convergence is the same as for $\lambda(0.4142)$. The values of $1/R^n a_n$ for the first nine terms of the series (considered as a power series in z^2 and given by (59) and (61)) have been calculated, and they are as follows:

term	2	3	4	5	6	7	8	9
$1/R^n a_n$	33.97	49.50	67.88	88.50	109.81	131.62	153.79	176.31

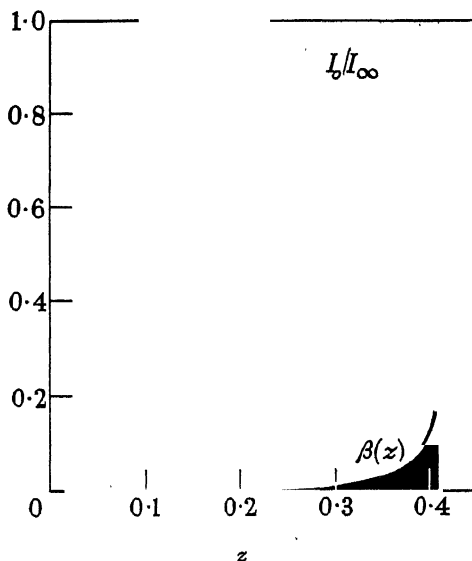
(62)


FIGURE 2

These terms are best considered logarithmically for the estimation of A and k . It is found that for all but the first few terms the values of k obtained decrease slowly, the value at the last term being about 1.16. When numerical values are substituted it seems likely that $\beta(x)$ is substantially different from 0.5, except in the immediate

neighbourhood of the singularity, where it comes up very rapidly to 0.5. The estimated $\beta(z)$ and the corresponding spontaneous magnetization I_0/I_∞ are plotted in figure 2.

7. THE HIGH-TEMPERATURE EXPANSIONS

Before proceeding further, we shall investigate the behaviour of the specific heat singularity in the presence of a magnetic field, and this involves an analysis of the series (50) for values of $\mu < 1$. We have tabulated the coefficients of powers of u ($= z^2$) in (50) for $\mu = 0.9, 0.8$ and 0.7 , and they are as follows:

μ	coefficient of					
	u^2	u^3	u^4	u^5	u^6	u^7
1	1	2	5	14	44	152
0.9	0.9	1.62	3.41	7.39	16.0	32.2
0.8	0.8	1.28	2.20	3.35	3.21	-5.48
0.7	0.7	0.98	1.32	1.10	-1.44	-10.9

(63)

It is not difficult to detect a fundamental difference in behaviour between terms for $\mu = 1$ and $\mu < 1$. When $\mu = 1$ the terms are all positive and steadily increase, and this corresponds to a singularity on the real axis (Dienes 1931). On the other hand, for $\mu = 0.7$ and 0.8 the terms do not remain consistent in sign; even for $\mu = 0.9$ the ratio of successive terms ceases to increase at the last term given above, which is not the case at any stage for $\mu = 1$; there are strong indications that for any $\mu < 1$ the terms are not consistent in sign beyond some point. This leads one to conjecture that the specific heat transition becomes continuous in the presence of a magnetic field, which is also the result given by the Heisenberg theory.

A rearrangement of the terms of the series (50) provides further evidence that there are no singularities in Λ for $\mu < 1$. Let us write Λ as a power series in μ whose coefficients are functions of z ; this corresponds to starting at some point of DE in figure 1 and moving parallel to DA . We obtain

$$\Lambda = 1 + \mu u^2 + \mu^2(2u^3 - 2u^4) + \mu^3(6u^4 - 14u^5 + 8u^6) + \mu^4(u^4 + 18u^5 - 77u^6 + 98u^7 - 40u^8) + \dots \quad (64)$$

This series is valid in the neighbourhood of DE for all u . At $u = 1$ it is easy to show that $\Lambda = 1 + \mu$. Hence all coefficients of powers of μ higher than the first must vanish when $u = 1$; it was, in fact, this consideration which enabled us to fill in the last term of (64) as $-40u^8$. (It is easy to see on physical grounds by considering overturned spins that the highest power of u in the term in μ^p is u^{2p} .)

We now try to find an expansion for values of u near $u = 1$. Putting $(1 - u) = \zeta$, and expanding Λ as a power series in ζ , the coefficient of ζ is

$$-2\mu + 2\mu^2 - 2\mu^3 + 2\mu^4 + \dots, \quad (65)$$

and one may assume that the sum of this is $-2\mu/(1 + \mu)$. Similarly the coefficient of ζ^2 is

$$\mu - 6\mu^2 + 16\mu^3 - 31\mu^4 + \dots, \quad (66)$$

and in this series second differences of coefficients are constant; this indicates that the sum is $\mu - 3\mu^2 + \mu^3/(1 + \mu)^3$. One is thus tempted to conjecture a general expansion of the form

$$\Lambda = 1 + \mu + \sum_{r=1}^{\infty} \frac{\phi_r(\mu) \zeta^r}{(1 + \mu)^{2r-1}}. \quad (67)$$

It should not be difficult to establish the form (67) rigorously by high-temperature expansions of the type referred to in § 1 (a). We can easily check the first two terms above by considering their values for $\mu = 1$, the transformation (3) being available here; it is found that they give the correct result.

Using the values of $\phi_1(\mu)$ and $\phi_2(\mu)$ and the series (50) we can now find the coefficient of μ^5 in (64); we can then return and determine $\phi_3(\mu)$. These are, respectively,

$$8u^5 + 44u^6 - 370u^7 + 799u^8 - 706u^9 + 225u^{10} \quad (68)$$

$$\text{and} \quad 6\mu^2 - 14\mu^3 + 6\mu^4. \quad (69)$$

This process can be extended if we notice a symmetry property of the $\phi_r(\mu)$. All the $\phi_r(\mu)$ derived so far satisfy the relation

$$\phi_r\left(\frac{1}{\mu}\right) = \frac{1}{\mu^{2r}} \phi_r(\mu). \quad (70)$$

This is exactly the relation we should expect from (41) if there were no singularity in Λ along BC in figure 1. It corresponds to the physical fact that there is no spontaneous magnetization in this region, as can be verified by evaluating (53). Assuming (70) we can deduce several more terms of (64) and (67); each additional term of (64) introduces two unknowns, and each term of (67) one. We find for (64),

$$\begin{aligned} & \mu^6(2u^5 + 40u^6 + 40u^7 - 1556u^8 + 5304u^9 - 7672u^{10} + 520u^{11} - 1362u^{12}) \\ & + \mu^7(22u^6 + 138u^7 - 424u^8 - 5470u^9 + 30348u^{10} \\ & - 65104u^{11} + 70850u^{12} - 39030u^{13} + 8670u^{14}), \end{aligned} \quad (71)$$

and for (67)

$$\left. \begin{aligned} \phi_4(\mu) &= -2\mu^2 + 42\mu^3 - 84\mu^4 + 42\mu^5 - 2\mu^6, \\ \phi_5(\mu) &= -34\mu^3 + 320\mu^4 - 580\mu^5 + 320\mu^6 - 34\mu^7, \\ \phi_6(\mu) &= 8\mu^3 - 423\mu^4 + 2591\mu^5 - 4365\mu^6 + 2591\mu^7 - 423\mu^8 + 8\mu^9, \\ \phi_7(\mu) &= 222\mu^4 - 4720\mu^5 + 21926\mu^6 - 34858\mu^7 + 21926\mu^8 - 4720\mu^9 + 222\mu^{10}. \end{aligned} \right\} \quad (72)$$

Again, the values of (72) at $\mu = 1$ have been checked and they are found to be correct. Finally, using the known values at $\mu = 1$ it is possible to add one more term to (71), and two more terms to (72).

All the results of this section have been deduced from the one series (50) transformed into (64). This brings out very clearly that the high- and low-temperature expansions for $\mu = 1$ are not independent, as one would expect from statistical considerations. The series (67) is valid for all μ ($\zeta < 0.828$); and the series (64) seems to be valid for all $\mu < 1$, thus strongly indicating the absence of any singularity in this region.

8. THE THEORY OF SOLID SOLUTIONS

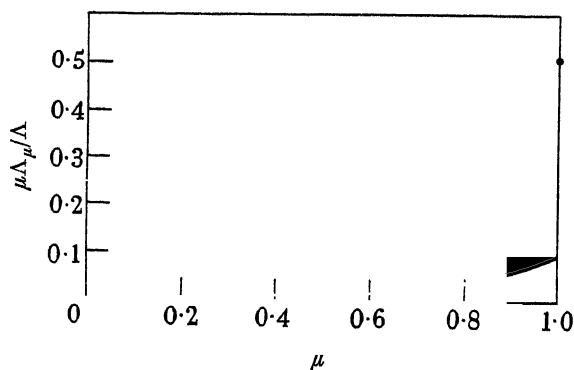
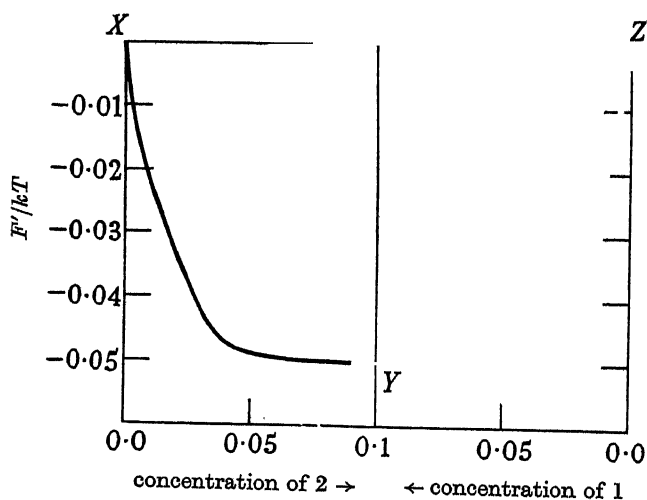
In I, § 2 the derivation of the partition function for the order-disorder and solid solution problems was discussed. For the quadratic lattice which we are considering, when the zero-energy terms are taken into account, the free energy per unit is given by

$$F' = -kT[\log \Lambda - \alpha_1 \log \mu] + 2\alpha_1 \epsilon'_{11} + 2(1 - \alpha_1) \epsilon'_{22}, \quad (73)$$

where

$$\mu \Lambda_\mu / \Lambda = \alpha_1. \quad (74)$$

When the sign of the energy $\epsilon'_{12} - \frac{1}{2}(\epsilon'_{11} + \epsilon'_{22})$ is considered, it will be found that values of z between 0 and 1, which we have been discussing in the present paper, correspond to the solid solution problem. The order-disorder problem corresponds to values of z between 0 and 1, and is thus related to the antiferromagnetic problem. The exact theory of binary solid solutions was considered by Lassettre & Howe (1941), and we shall find many of their conclusions here substantiated.

FIGURE 3. ($z = 0.4$.)FIGURE 4. ($z = 0.4$.)

If we confine our attention to a fixed concentration α_1 , the relation (74) determines a path in the (μ, z) plane along which Λ must be evaluated to determine F' . When $\alpha_1 = \frac{1}{2}$ the path is AB . When α_1 is different from $\frac{1}{2}$ a different type of path results. When $z < 0.414$ the general behaviour of $\mu\Lambda_\mu/\Lambda$ as a function of μ is similar to the curve in figure 3 ($z = 0.4$). Its value increases steadily from zero at $\mu = 0$ to $\beta(z)$ at $\mu = 1$ and there is a jump to $\frac{1}{2}$ at $\mu = 1$. This discontinuity arises for an infinite lattice, and would be a very rapid continuous change for a large finite lattice. Hence if $\beta(z) < \alpha_1 < \frac{1}{2}$ the relation (74) can only be satisfied on the line AB . For $z = 1$ we can use (67) to show that $\mu/1 + \mu \rightarrow \alpha_1$. Thus for any value of α_1 different from $\frac{1}{2}$ the

path given by (74) moves along AC until $\beta(z) = \alpha_1$, and then it turns into the plane cutting BE in a point $\mu = \alpha_1/1 - \alpha_1$. Typical paths are AC_1B_1 , AC_2B_2 .

Clearly a transition takes place at points such as C_1 , and we shall interpret the portion AC_1 as corresponding to a separation into two phases, and the portion C_1B as solution. This interpretation is borne out by the curve of free energy as a function of concentration shown in figure 4. (For simplicity the case $\epsilon'_{11} = \epsilon'_{22}$ has been taken, but the resulting features are essentially the same for the general case.) The portion AC_1 may be regarded as a mixture of two phases at Y , phase I being a solution of $\beta(z)$ of substance 1 in substance 2, and phase II a solution of $\beta(z)$ of substance 2 in substance 1. $\beta(z)$ is thus the maximum concentration which can exist in solution at the given temperature, and is in fact the solubility curve for the two substances; it is shown in figure 2. As we move along AC_1 the proportion of phase II gets smaller until at C_1 it disappears.

The nature of the transition at C_1 can be determined from (73) and (74). There is a discontinuity in μ_z at C_1 given by

$$\lim_{\mu \rightarrow 1} \frac{\alpha_1 \Lambda_z - \mu \Lambda_{\mu z}}{\mu \Lambda_{\mu\mu} + (1 - \alpha_1) \Lambda_{\mu}}. \quad (75)$$

On the other hand, $(F'/T)_{\mu} = 0$, so that the energy E is equal to $\epsilon_{12} z \Lambda z / \Lambda$ and this does not involve μ_z . Thus the energy is continuous, but the specific heat has a discontinuity given by

$$\left. \begin{aligned} \left[\frac{\epsilon_{12}^2}{kT^2} z^2 \mu_z \frac{\partial}{\partial \mu} (\Lambda z / \Lambda) \right]_{\mu \rightarrow 1} &= k(z \log z)^2, \\ \lim_{\mu \rightarrow 1} \left[\mu_z \frac{\partial}{\partial \mu} (\Lambda z / \Lambda) \right] &, \end{aligned} \right\} \quad (76)$$

so that the transition is second order. For $\alpha_1 = 0.5$ the transition takes place at C , and the specific heat curve has a logarithmic infinity of the type given by Onsager (1944).

Finally, it seems reasonable to assume that with the exception of the detailed behaviour at the singularity C , the discussion of § 8 should be qualitatively applicable to a three-dimensional model.

The work described above was carried out at Cambridge during the tenure of a Senior Research Award of the D.S.I.R.

REFERENCES

- Aitken, A. C. 1937 *Proc. Roy. Soc. Edinb.* **57**, 269.
 Ashkin, J. & Lamb, W. E. 1943 *Phys. Rev.* **64**, 159.
 Dienes, P. 1931 *The Taylor series*, chapter xiv. Oxford: University Press.
 Domb, C. 1949 *Proc. Roy. Soc. A*, **196**, 36. (Referred to as I.)
 Kramers, H. A. & Wannier, G. H. 1941 *Phys. Rev.* **60**, 252, 263.
 Lassettre, E. N. & Howe, J. P. 1941 *J. Chem. Phys.* **9**, 747, 781.
 Onsager, L. 1944 *Phys. Rev.* **65**, 117.
 Wannier, G. H. 1945 *Rev. Mod. Phys.* **17**, 50.

For table 1 see pull-out sheet facing p. 220.

TABLE 2

x_1							$\lambda_2 = 0$
x_2	$0_r 1$						$\lambda_3 = 0$
x_3	10_r	$0_r 1_2$					$\lambda_4 = 1$
	$0_r 10_s$						
x_4	$0_r 1$	$1_2 0_r$	$0_r 1_3$				$\lambda_5 = 0$
		$0_r 1_2 0_s$					
x_5	$10_r = 2$	$0_r 1_2 = 3$	$1_3 0_r$	$0_r 1_4$			$\lambda_6 = 2$
	$0_r 10_s$	$10_r 1 = 2$	$0_r 1_3 0_s$				
		$0_r 10_s 1$					
x_6	$0_r 1$	$1_2 0_r = 4$	$0_r 1_3 = 6$	$1_4 0_r$	$0_r 1_5$		$\lambda_7 = 0$
		$0_r 1_2 0_s = 3$	$10_r 1_2 = 3$	$0_r 1_4 0_s$			
		$10_r 10_s$	$1_2 0_r 1 = 2$				
		$0_r 10_s 10_t$	$0_r 1_a 0_s 1_b^*$				
x_7	$10_r = 5$	$0_r 1_2 = 5$	$1_3 0_r = 7$	$0_r 1_4 = 10$	$1_5 0_r$	$0_r 1_6$	$\lambda_8 = 5$
	$0_r 10_s = 2$	$10_r 1 = 5$	$0_r 1_3 0_s = 6$	$10_r 1_3 = 4$	$0_r 1_5 0_s$		
		$0_r 10_s 1 = 2$	$1_a 0_r 1_b 0_s$	$1_2 0_r 1_2 = 3$			
			$0_r 1_a 0_s 1_b 0_t$	$1_3 0_r 1 = 2$			
				$0_r 1_a 0_s 1_b$			
x_8	$0_r 1 = 2$	$1_2 0_r = 13$	$0_r 1_3 = 16$	$1_4 0_r = 11$	$0_r 1_5 = 15$	$1_6 0_r$	$\lambda_9 = 0$
		$0_r 1_2 0_s = 7$	$10_r 1_2 = 12$	$0_r 1_4 0_s = 10$	$10_r 1_4 = 5$	$0_r 1_6 0_s$	
		$10_r 10_s = 3$	$1_2 0_r 1 = 10$	$1_a 0_r 1_b 0_s$	$1_3 0_r 1_3 = 4$		
		$0_r 10_s 10_t = 2$	$0_r 1_a 0_s 1_b = 4$	$0_r 1_a 0_s 1_b 0_t$	$1_3 0_r 1_2 = 3$		
		$10_r 10$	$10_r 10_s 1 = 2$		$1_4 0_r 1 = 2$		
			$0_r 10_s 10_t 1$		$0_r 1_a 0_s 1_b$		

* $0_r 1_a 0_s 1_b$ refers to all configurations of this type such that $a > 0$, $b > 0$, $a + b$ = number of 1's in the column, and similarly for other configurations.

TABLE 3

v_1	10_r				$\lambda_3 = 0$
	$0_r 10_s$				
v_2	$0_r 1$	$1_2 0_r$			$\lambda_4 = 1$
		$0_r 1_2 0_s$			
		$10_r 10_s$			
		$0_r 10_s 10_t$			
v_3	10_r	$0_r 1_2$	$1_3 0_r$		$\lambda_5 = 0$
	$0_r 10_s$	$10_r 1 = 2$	$0_r 1_3 0_s$		
		$0_r 10_s 1$	$1_a 0_r 1_b 0_s$		
			$0_r 1_a 0_s 1_b 0_t$		
			$10_r 10_s 10_t$		
			$0_r 10_s 10_t 10_u$		
v_4	$0_r 1 = 2$	$1_2 0_r = 3$	$0_r 1_3$	$1_4 0_r$	$\lambda_6 = 2$
		$0_r 1_2 0_s = 3$	$10_r 1_2 = 2$	$0_r 1_4 0_s$	
		$10_r 10_s = 2$	$1_2 0_r 1 = 3$	$1_a 0_r 1_b 0_s$	
		$0_r 10_s 10_t = 2$	$0_r 1_a 0_s 1_b$	$0_r 1_a 0_s 1_b 0_t$	
			$10_r 10_s 1 = 2$	$1_a 0_r 1_b 0_s 1_c 0_t$	
			$0_r 10_s 10_t 1$	$0_r 1_a 0_s 1_b 0_t 1_c 0_u$	
				$10_r 10_s 10_t 10_u$	
				$0_r 10_s 10_t 10_u 10_v$	

TABLE 4

X_{11} $0, 1$	X_{12} $0, 1_2$	X_{13} $0, 1_3$	X_{14} $0, 1_4$	X_{15} $0, 1_5$	X_{16} $0, 1_6$
X_{21} $10, 1$ $0, 10_2$ $\lambda_{22} = 1$	X_{22} $1_2 0, 1$ $0, 1_2 0_2$	X_{23} $10, 1$ $\lambda_{24} = 1$	X_{24} $1_2 0, 1$ $1_4 0, 1$ $0, 1_4 0_2$	X_{25} $10, 1$ $1_3 0, 1$ $\lambda_{26} = 1$	X_{26} $1_2 0, 1$ $1_4 0, 1$ $1_6 0, 1$ $0, 1_6 0_2$
X_{31} $0, 1$	X_{32} $0, 1_2 = 3$ $10, 1 = 2$ $0, 10_2 1$	X_{33} $0, 1_3 = 6$ $10, 1_2 = 3$ $1_2 0, 1 = 2$	X_{34} $10, 1$ $0, 10_2 1_3$ $0, 1_2 0, 1_3$ $0, 1_3 0, 1$	X_{35} $1_2 0, 1$ $10, 1_2$	X_{36} $10, 1$ $10, 1_5 = 6$ $1_2 0, 1_4 = 5$ $1_3 0, 1_3 = 4$ $1_3 0, 1_2 = 3$ $1_4 0, 1 = 2$ $0, 1_4 0_2 1_3$
X_{41} $10, 1$ $0, 10_2$ $\lambda_{43} = 1$	X_{42} $1_2 0, 1 = 3$ $0, 1_2 0_2 = 3$ $10, 10_2$ $0, 10_2 10_2$	X_{43} $10, 1 = 3$ $0, 10_2$	X_{44} $1_2 0, 1 = 7$ $0, 1_2 0_2 = 2$ $10, 10_2$ $10, 10_2$ $0, 1_2 0_2 1_3 0_2$	X_{45} $10, 1 = 6$ $0, 10_2 0$ $1_5 0, 1 = 15$ $0, 1_5 0_2 = 15$ $1_6 0, 1_5 0_2$ $\lambda_{46} = 6$	X_{46} $1_2 0, 1 = 12$ $0, 1_2 0_2 = 2$ $10, 10_2$ $10, 10_2$ $1_6 0, 1 = 21$ $0, 1_6 0_2 = 21$ $1_4 0, 1_5 0_2$ $0, 1_4 0_2 1_5 0_2$

	x_0	x_1	$A_1 x_1$	$A_2 x_0$	u_2	x_2	$A_1 x_2$	$A_2 x_1$	u_3	$A_0 u_3$	$A_0^2 u_3$	$A_0^3 u_3$	x_3	$A_1 x_3$	$A_2 x_2$	u_2	$A_0 u_4$	$A_0^2 u_4$	$A_0^3 u_4$	x_4
00000(11111)	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
00001(11110)	.	.	.	1	1	1	.	.	.	.	.	.	.	1	.	1	.	.	.	1
00010(11101)	.	.	.	.	.	.	1	.	1	.	.	.	1	.	.	.	.	.	.	.
00011(11100)	.	.	.	.	.	.	1	.	1	.	.	.	1	.	.	.	.	.	.	.
00100(11011)	.	.	.	.	.	.	.	.	.	1	.	.	1	.	.	.	.	.	.	.
00101(11010)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
00110(11001)	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	1	.	.	.	1
00111(11000)	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	1	.	1	.	2
01000(10111)	.	.	.	.	.	.	.	.	.	.	1	.	1	.	.	.	.	.	.	.
01001(10110)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
01010(10101)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
01011(10100)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
01100(10011)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	1
01101(10010)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
01110(10001)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
01111(10000)	.	.	.	.	.	.	.	.	.	.	.	1	1	.	.	.	.	.	.	.

	y_0	$A'_1 y_0$	v_1	$A'_0 v_1$	$A_0'^2 v_1$	$A_0'^3 v_1$	y_1	$A'_1 y_1$	$A'_2 y_0$	v_2	$A'_0 v_2$	$A_0'^2 v_2$	$A_0'^3 v_2$	y_2	$A'_1 y_2$	$A'_2 y_1$	v_3	$A'_2 v_3$	$A_0'^3 v_3$	$A_0'^3 v_3$
00000(11111)	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
00001(11110)	.	.	.	.	.	.	.	1	.	1	.	.	.	1	.	.	.	.	.	.
00010(11101)	.	.	.	.	.	1	1	.	.	.	.	.	.	.	.	.	.	.	.	1
00011(11100)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2	.	2	.	.	.
00100(11011)	.	.	.	.	1	.	1	.	.	.	.	.	.	.	.	.	.	.	1	.
00101(11010)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2	.	2	.	.	.
00110(11001)	.	.	.	.	.	.	.	.	.	.	1	.	1	.	.	.	.	.	.	.
00111(11000)	.	.	.	.	.	.	.	1	.	1	.	.	.	1	.	.	.	.	.	.
01000(10111)	.	.	.	1	.	.	1	.	.	.	.	.	.	.	.	.	.	1	.	.
01001(10110)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2	.	2	.	.	.
01010(10101)	.	.	.	.	.	.	.	.	.	.	1	.	.	1	.	.	.	.	.	.
01011(10100)	.	.	.	.	.	.	.	1	.	1	.	.	.	1	.	.	.	.	.	.
01100(10011)	.	.	.	.	.	.	.	.	.	.	1	.	.	1	.	.	.	.	.	.
01101(10010)	.	.	.	.	.	.	.	1	.	1	.	.	.	1	.	.	.	.	.	.
01110(10001)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1	2	.	.
01111(10000)	.	1	1	.	.	.	1	.	.	.	.	.	.	.	1	.	1	.	.	.

TABLE 1

A_1x_4	A_2x_3	u_5	A_0u_5	$A_0^2u_5$	$A_0^3u_5$	x_5	A_1x_5	A_2x_4	$-x_2$	u_6	A_0u_6	$A_0^2u_6$	$A_0^3u_6$	x_6	A_1x_6	A_2x_5	$-x_3$	u_7	A_0u_7	$A_0^2u_7$	$A_0^3u_7$	x_7
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1	.	1	.	.	.	1	3	.	-1	2	.	.	.	2	2	3	-1	4	.	.	.	4
1	.	1	3	.	.	4	.	.	.	.	.	.	.	.	2	.	-1	1	11	.	.	12
.	.	.	1	.	.	1	.	.	.	.	.	.	.	.	2	.	-1	1	4	.	.	5
.	1	1	.	.	.	1	.	.	.	.	.	.	.	.	2	1	.	3	.	.	.	3
.	.	.	.	.	.	.	4	1	.	5	.	.	.	5	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	4	.	.	4	1	5	.	10	.	.	.	.	.	.	.	.
.	.	.	.	1	.	1	.	.	.	.	.	.	.	.	2	.	-1	1	1	4	.	6
.	1	1	.	.	.	1	.	.	.	.	.	.	.	.	2	1	.	3	.	.	.	3
.	.	.	.	.	.	.	1	.	.	1	.	.	.	1	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	1	.	.	1	1	.	.	2	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	1	.	.	1	5	.	.	6	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	1	1	.	2	.	.	.	2	.	.	.	.	.	.	.	.
2	1	3	.	.	.	3	.	.	.	.	.	.	.	.	10	1	.	11	.	.	.	11
2	.	2	.	.	1	3	.	.	.	.	.	.	.	.	10	.	-1	9	1	1	4	15

y_3	$A_1'y_3$	$A_2'y_2$	v_4	$A_0'v_4$	$A_0'^2v_4$	$A_0'^3v_4$	y_4	$A_1'y_4$	$A_2'y_3$	$-y_1$	v_5	$A_0'v_5$	$A_0'^2v_5$	$A_0'^3v_5$	y_5	$A_1'y_5$	$A_2'y_4$	$-y_2$	v_6	$A_0'v_6$	$A_0'^2v_6$	$A_0'^3v_6$	y_6
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	3	.	3	.	.	.	3	.	.	.	.	.	.	.	.	16	.	-1	15	.	.	.	15
1	.	.	.	.	.	.	.	.	2	-1	1	1	2	2	6	.	.	.	.	.	.	.	.
2	.	.	.	.	.	.	.	10	.	.	10	.	.	.	10	.	.	.	.	.	.	.	.
1	.	.	.	.	.	.	.	.	2	-1	1	2	2	.	5	.	.	.	.	.	.	.	.
2	.	.	.	.	.	.	.	7	.	.	7	.	.	.	7	.	.	.	.	.	.	.	.
.	.	1	1	1	4	.	6	.	.	.	.	.	.	.	.	.	3	-1	2	5	12	.	19
.	4	.	4	.	.	.	4	.	.	.	.	.	.	.	.	13	.	-1	12	.	.	.	12
1	.	.	.	.	.	.	.	.	3	-1	2	2	.	.	4	.	.	.	.	.	.	.	.
2	.	.	.	.	.	.	.	8	.	.	8	.	.	.	8	.	.	.	.	.	.	.	.
.	.	1	1	3	.	.	4	.	.	.	.	.	.	.	.	.	4	-1	3	11	.	.	14
.	3	.	3	.	.	.	3	.	.	.	.	.	.	.	.	12	.	-1	11	.	.	.	11
.	.	1	1	4	.	.	5	.	.	.	.	.	.	.	.	.	6	-1	5	12	.	.	17
.	3	.	3	.	.	.	3	.	.	.	.	.	.	.	.	12	.	-1	11	.	.	.	11
3	.	.	.	.	.	.	.	.	1	.	1	10	.	.	11	.	.	.	.	.	.	.	.
1	.	.	.	.	.	.	.	3	.	-1	2	.	.	.	2	.	.	.	.	.	.	.	.

TABLE 7

z_1	$10_r = \mu z^2/W_1$	$0_r 10_s = \mu z^2/W_1$	$0_r 1 = \mu z/W_1$	
z_2	$1_2 0_r = (\mu^2 z^2 + \mu^3 z^4)/W_1 W_2$	$10_r 10_s = \mu^2 z^4/W_1^2$	$0_r 10_s 10_t = \mu^2 z^4/W_1^2$	$W_r = 1/(1 - \mu^2 z^2)$
	$0_r 1_2 0_s = (\mu^2 z^2 + \mu^3 z^4)/W_1 W_2$	$0_r 10_s 1 = \mu^2 z^3/W_1^2$		
	$0_r 1_2 = (\mu^2 z + \mu^3 z^3)/W_1 W_2$			
	$10_r 1 = (\mu^2 + \mu^3) z^3/W_1 W_2$			
z_3	$10_r = \frac{\mu^2 z^2 + (\mu^2 + 2\mu^3) z^4}{W_1^2 W_2} - \frac{3\mu^2 z^4}{W_1^3}$	$1_3 0_r = \frac{\mu^4 z^4}{W_1^2 W_3} + \frac{\mu^3 z^2 + (\mu^4 + \mu^5) z^4 + \mu^6 z^6}{W_1 W_2 W_3}$	$0_r 1_3 0_s = \frac{\mu^4 z^3}{W_1^2 W_3} + \frac{\mu^3 z + (\mu^4 + \mu^5) z^3 + \mu^6 z^5}{W_1 W_2 W_3}$	
	$0_r 10_s = \frac{(\mu^2 + 2\mu^3) z^4 + \mu^4 z^6}{W_1^2 W_2} - \frac{2\mu^2 z^4 + \mu^3 z^6}{W_1^3}$	$0_r 1_3 = \frac{\mu^4 z^3}{W_1^2 W_3} + \frac{\mu^3 z + (\mu^4 + \mu^5) z^3 + \mu^6 z^5}{W_1 W_2 W_3}$	$10_r 1_2 = \frac{\mu^2 z^3}{W_1^2 W_3} + \frac{(\mu^4 + \mu^5) z^3 + (\mu^5 + \mu^6) z^5}{W_1 W_2 W_3}$	
	$0_r 1 = \frac{\mu^3 z^3 + (\mu^3 + 2\mu^4) z^5}{W_1^2 W_2} - \frac{\mu^2 z^3 + 2\mu^3 z^5}{W_1^3}$	$1_2 0_r 1 = \frac{\mu^5 z^5}{W_1^2 W_3} + \frac{(\mu^3 + \mu^4) z^3 + (\mu^4 + \mu^5) z^5}{W_1 W_2 W_3}$		
	$\left. \begin{array}{l} 10_r 1_2 0_s \\ 1_2 0_r 10_s \\ 0_r 10_s 1_2 0_t \\ 0_r 1_2 0_s 10_t \end{array} \right\} = \frac{\mu^3 z^4 + \mu^4 z^6}{W_1^2 W_2}$	$\left. \begin{array}{l} 0_r 10_s 1_2 \\ 0_r 1_2 0_s 1 \\ 10_r 10_s 1 \end{array} \right\} = \frac{\mu^3 z^3 + \mu^4 z^5}{W_1^2 W_2}$	$0_r 10_s 10_t 1 = \mu^3 z^5/W_1^3$	$\left. \begin{array}{l} 10_r 10_s 10_t \\ 0_r 10_s 10_t 10_u \end{array} \right\} = \mu^3 z^6/W_1^3$
z_4	$1_2 0_r = \frac{\mu^3 z^2 + 2(\mu^4 + \mu^5) z^4 + (\mu^5 + 2\mu^6) z^6}{W_1 W_2^2 W_3} + \frac{(\mu^3 + \mu^4) z^4}{W_1^2 W_2 W_3}$			
	$+ \frac{(-3\mu^3 + \mu^4) z^4 + (-2\mu^4 + 3\mu^5) z^6 + \mu^6 z^8}{W_1^2 W_2^2} - \frac{\mu^2 z^4 + 4\mu^4 z^6 + \mu^5 z^8}{W_1^3 W_2}$			
	$0_r 1_2 0_s = \frac{(\mu^4 + 2\mu^5) z^4 + (\mu^5 + 2\mu^6 + \mu^7) z^6 + \mu^8 z^8}{W_1 W_2^2 W_3} + \frac{\mu^3 z^4 + \mu^6 z^6}{W_1^2 W_2 W_3}$			
	$+ \frac{(-2\mu^3 + \mu^4) z^4 + (-\mu^4 + 2\mu^5) z^6}{W_1^2 W_2^2} - \frac{\mu^3 z^4 + 4\mu^4 z^6 + \mu^5 z^8}{W_1^3 W_2}$			
	$0_r 1_2 = \frac{\mu^5 z^3 + 2(\mu^6 + \mu^7) z^5 + (\mu^6 + 2\mu^7) z^7}{W_1 W_2^2 W_3} + \frac{(\mu^5 + \mu^6) z^5}{W_1^2 W_2 W_3}$			
	$+ \frac{(-\mu^3 + \mu^4) z^3 + \mu^5 z^5 - \mu^6 z^7}{W_1^2 W_2^2} - \frac{\mu^3 z^3 + 4\mu^4 z^5 + \mu^5 z^7}{W_1^3 W_2}$			
	$10_r 1 = \frac{\mu^4 z^3 + 2(\mu^5 + \mu^6) z^5 + (\mu^6 + 2\mu^7) z^7}{W_1 W_2^2 W_3} + \frac{(\mu^4 + \mu^5) z^5}{W_1^2 W_2 W_3}$			
	$+ \frac{(-2\mu^4 + \mu^5) z^5 + (-\mu^5 + 2\mu^6) z^7}{W_1^2 W_2^2} - \frac{(2\mu^3 + \mu^4) z^5 + (\mu^4 + 2\mu^5) z^7}{W_1^3 W_2}$			
	$10_r 10_s = \frac{\mu^3 z^4 + (2\mu^3 + 4\mu^4) z^6 + \mu^5 z^8}{W_1^2 W_2} - \frac{5\mu^3 z^6 + \mu^4 z^8}{W_1^3}$			
	$0_r 10_s 10_t = \frac{(2\mu^3 + 4\mu^4) z^6 + 2\mu^5 z^8}{W_1^2 W_2} - \frac{4\mu^3 z^6 + 2\mu^4 z^8}{W_1^3}$			
	$0_r 10_s 1 = \frac{(\mu^3 + 3\mu^4) z^5 + (\mu^4 + 3\mu^5) z^7}{W_1^2 W_2} - \frac{3\mu^3 z^5 + 3\mu^4 z^7}{W_1^3}$			
	$\left. \begin{array}{l} 10_r 10 \\ 0_r 10 1 \\ 0_r 10 10_s \\ 10 10_r \\ 0 10_r 1 \end{array} \right\} = \frac{1}{1 - \mu^2 z^4} \left\{ \frac{\mu^5 z^6}{W_1^2 W_3} + \frac{(\mu^3 + \mu^4) z^4 + (\mu^4 + \mu^5) z^6}{W_1 W_2 W_3} - \frac{\mu^3 z^4 + (\mu^3 + 2\mu^4) z^6}{W_1^2 W_2} + \frac{\mu^2 z^6}{W_1^3} \right\}$			
z_5	$10_r = \frac{\mu^3 z^2 + (3\mu^4 + 2\mu^5) z^4 + (3\mu^5 + 4\mu^6) z^6 + (\mu^6 + 2\mu^7) z^8}{W_1^2 W_2^2 W_3} + \frac{2(\mu^3 + \mu^4) z^4 + 2(\mu^4 + \mu^5) z^6}{W_1^3 W_2 W_3}$			
	$+ \frac{\mu^5 z^6}{W_1^4 W_3} + \frac{(-3\mu^3 + \mu^4) z^4 + (-4\mu^4 + 4\mu^5) z^6 + (-\mu^5 + 3\mu^6) z^8}{W_1^3 W_2^2}$			
	$- \frac{6\mu^3 z^4 + (8\mu^3 + 18\mu^4) z^6 + (2\mu^4 + 6\mu^5) z^8}{W_1^4 W_2} + \frac{16\mu^3 z^6 + 3\mu^4 z^8}{W_1^5}$			

Nuclear magnetic resonance at low temperatures

BY J. HATTON AND B. V. ROLLIN

(Communicated by F. Simon, F.R.S.—Received 22 April 1949)

Investigations have been made on nuclear paramagnetism at low temperatures by the radio-frequency resonance absorption method. In measurements on solid hydrogen it has been found that there are considerable changes in the width of the resonance line with change of temperature. These effects are related to changes in the nature of the molecular motions and the state of the molecular order. A number of experiments have been made on nuclear spin-lattice relaxation in different substances. It has been shown that in ionic crystals the nuclear relaxation is greatly affected by the presence of paramagnetic impurities. The nuclear relaxation time in metals has been found to be in reasonable agreement with theoretical predictions. Measurements have been extended to temperatures below 1°K , and the possibility of cooling the nuclear spin system by adiabatic demagnetization has been verified. The results for nuclear spin-lattice relaxation are considered in relation to the possible applications of nuclear paramagnetism in the region of very low temperatures.

1. INTRODUCTION

If a substance containing N nuclei with magnetic moment μ and spin $I\hbar/2\pi$ is placed in a magnetic field H_0 , a magnetization M_0 is set up, given by the Curie relation

$$M_0 = \chi_0 H_0 = N \frac{\mu^2}{3kT} \frac{I+1}{I} H_0. \quad (1)$$

This magnetization is attained only when thermal equilibrium has been established, so that there is a Boltzmann distribution, corresponding to the lattice temperature T , over the $2I+1$ magnetic energy levels. The rate at which equilibrium is attained depends on the rate at which energy and angular momentum can be transferred from the nuclei to the lattice, and the approach to equilibrium may be described by the relaxation equation

$$\frac{dM}{dt} = \frac{M_0 - M}{T_1}, \quad (2)$$

where T_1 is the spin-lattice relaxation time.

The relaxation time T_1 can be measured by radio-frequency resonance methods (Purcell, Torrey & Pound 1946; Rollin 1946; Bloch, Hansen & Packard 1946). The essential feature of these methods is that the substance is subjected to a strong magnetic field H_0 , and a weak oscillating field is applied in a transverse direction. The frequency ν_0 at which resonance is obtained is given by

$$h\nu_0 = \frac{\mu}{I} H_0. \quad (3)$$

For a value of H_0 of a few thousand gauss, the frequency required is of the order of 1 to 10 Mc./sec. There are a number of ways in which the resonance can be detected; one of the most convenient is that due to Purcell *et al.*, in which the resonance absorption of energy by the nuclei is measured by the damping of the tuned circuit which supplies the radio-frequency field. By finding the dependence of the absorp-

tion on the strength of the radio-frequency field, the spin-lattice relaxation time can be obtained.

The theory of spin-lattice relaxation was originally developed by Waller (1932) for the electron spins in paramagnetic substances. If this theory is applied to the nuclear spins in a non-metallic solid, it leads to nuclear spin-lattice relaxation times of the order of 10^4 sec. at room temperature and 10^{17} sec. at 1° K. If the relaxation time were indeed as long as this, it would preclude the possibility of making any practical application of nuclear paramagnetism such as the attainment of very low temperatures by adiabatic nuclear demagnetization (Gorter 1934; Kurti & Simon 1935) or the use of the nuclear susceptibility for measurement of temperatures near absolute zero (Rollin, Hatton, Cooke & Benzie 1947).

The relaxation time for the nuclei in ionic crystals has been found to differ considerably from the values predicted by the theory (Rollin 1946; Rollin & Hatton 1947; Rollin *et al.* 1947; Rollin & Hatton 1948). The disagreement is most pronounced at liquid helium temperatures, the measured values of relaxation time being only of the order of a few seconds or minutes. It appears that this discrepancy is due to the presence in the specimens of minute amounts of paramagnetic impurity which exert a considerable influence on the relaxation process.

One of the most interesting applications of the resonance method, apart from the study of nuclear relaxation, is to the investigation of the mutual interactions of the nuclei in a substance. The magnitude of the spin-spin interaction is indicated by the width of the absorption line. During the course of experiments on liquid and solid hydrogen, considerable changes in the width and shape of the line with change of temperature have been observed (Rollin & Hatton 1948). These changes can be correlated with changes in the molecular motions and state of order as the temperature is lowered.

2. EXPERIMENTAL METHOD

As the theory and methods of nuclear resonance have been extensively described in recent publications (Bloch 1946; Bloembergen, Purcell & Pound 1948; Bloembergen 1948), we will discuss only those features of our apparatus and techniques which differ from those of other workers.

The circuit we have employed is shown in figure 1. The specimen is placed in the coil of the tuned circuit which is supplied with radio-frequency current at the resonance frequency. The current is provided by a signal generator feeding through a high impedance so that we have essentially a constant current arrangement, and the radio-frequency voltage across the circuit is proportional to the circuit impedance. When the field H_0 is adjusted to the correct value for nuclear resonance, there is a drop in circuit impedance, on account of the nuclear absorption, and the resonance is observed as a change in radio-frequency voltage across the circuit. In order to facilitate observation, the field is modulated at an audio-frequency so that the change of nuclear absorption gives rise to an audio-frequency modulation of the radio-frequency voltage across the circuit. After amplification and rectification, the resulting audio-frequency signal can be observed on a cathode-ray oscillograph or measured by a meter.

As compared with the bridge balancing method used by Purcell *et al.* (1946), this arrangement has the advantage of simplicity and also freedom from perturbing effects due to phase changes arising from frequency modulation of the generator or vibration of the circuit elements. The principal disadvantage of our method is that rather stringent requirements are placed on the radio-frequency amplifier which must have variable gain, be linear and have a good noise factor over a wide range of input signals. To fulfil these requirements it is desirable to use an amplifier with negative feedback. Some form of balancing method would be necessary if a radio-frequency amplitude considerably greater than 1 V were required across the tuned circuit, in order to avoid saturation of the amplifier. Fortunately, all our measurements could be made with an amplitude less than 1 V. The smallest signal which could be detected was determined by the Johnson noise of the tuned circuit when the radio-frequency amplitude was less than 10 mV, and by microphonics for higher amplitudes. The microphonics were especially large on account of the difficulty of obtaining complete rigidity of construction of the coil and the leads to the cryostat in which the specimen was situated.

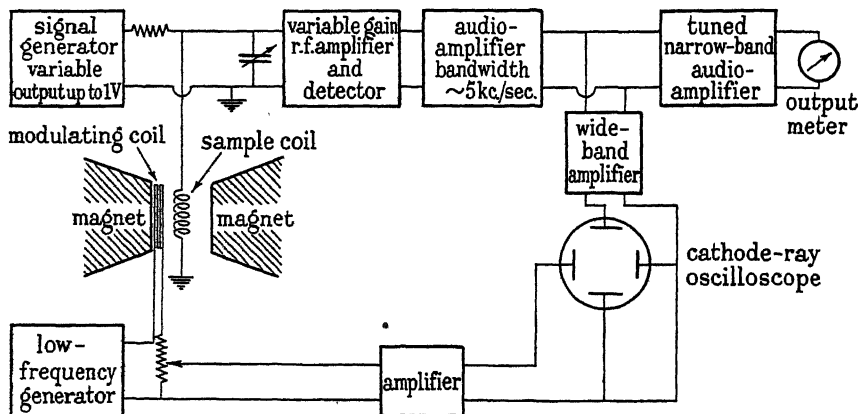


FIGURE 1

The field H_0 was provided by a small electromagnet capable of giving up to 4000 gauss with a 2 in. pole-gap. Most of the measurements were made at frequencies in the neighbourhood of 2 or 5 Mc./sec.

The specimen was generally of cylindrical shape, with a volume of about 1 cm.³, although much smaller specimens have been used. The specimen was contained in a glass cryostat attached to a helium liquefier of the Simon expansion type. The cryostat consisted of two vessels *A* and *B* situated in an evacuated enclosure (figure 2). The coil of the radio-frequency circuit was wound on the outside of vessel *A* which contained the specimen. Vessel *A* was provided with a ground joint so that the specimen could easily be changed. This joint was sealed with glycerine and was found to be perfectly vacuum-tight even when *A* contained liquid helium at temperatures below the λ -point.

For measurements at temperatures below 1°K, the specimen was cooled by adiabatic demagnetization of a paramagnetic salt. These experiments presented

considerable difficulty, as it was necessary to separate the specimen and the salt by an appreciable distance, in order to avoid excessive influence of H_0 on the salt, and at the same time it was necessary to maintain good thermal contact between them. After a number of experiments, it was found that the most satisfactory arrangement was to enclose the specimen and the salt in a glass capsule which was filled with helium gas at atmospheric pressure and room temperature and sealed off. The capsule was placed in vessel *A* and was cooled to approximately 1°K by liquid helium in *B*, the thermal contact being established by means of helium exchange gas in *A*. Before demagnetization, the exchange gas was pumped away so that the capsule was thermally insulated.

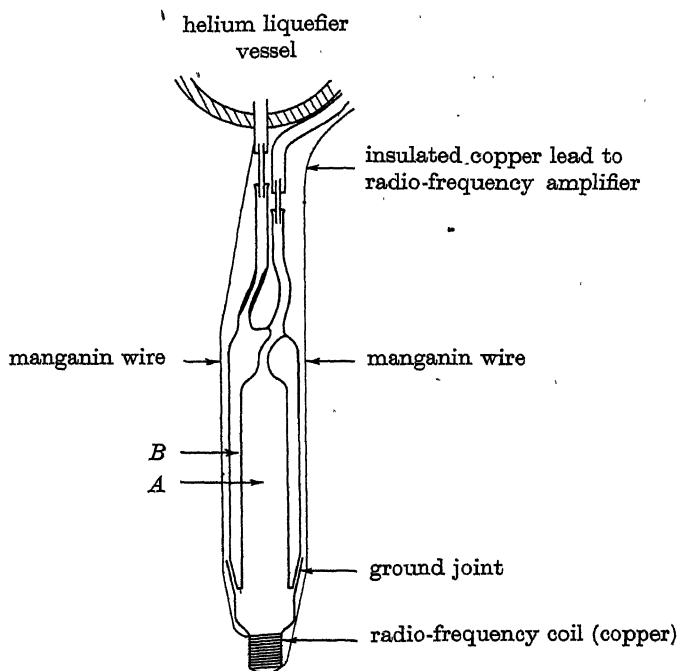


FIGURE 2

3. METHODS OF MEASURING THE SPIN-LATTICE RELAXATION TIME

The rate of absorption of energy by the nuclei is given by

$$P = \pi \nu H_1^2 \chi'',$$

where H_1 is the amplitude of radio-frequency field and χ'' , the absorption coefficient, is a function of ν and H_1 . If H_1 is very small, the maximum value of χ'' is given by

$$\chi'' = \frac{\nu_0}{\Delta\nu} \chi_0, \quad (4)$$

where $\Delta\nu$ is the width of the resonance. The change in voltage across the circuit is proportional to $\nu = \chi'' H_1$.

When H_1 is large, the absorption of energy by the nuclear-spin system causes an elevation of the spin temperature and a reduction of χ'' . The extent of this reduction

depends both on the magnitude of H_1 and on the spin-lattice interaction, so that by finding the value of H_1 required to cause an appreciable reduction in χ'' , the spin-lattice relaxation time can be calculated. It can be shown (Bloch 1946; Bloembergen *et al.* 1948; Bloembergen 1948) that ν has a maximum value when

$$\gamma^2 H_1^2 T_1 / \Delta \nu \simeq 1, \quad (5)$$

where $\gamma = 2\pi\mu/Ih$.

If T_1 is longer than a few seconds, it is more convenient to employ a direct method of measurement. A very strong radio-frequency field is applied so that the nuclear-spin temperature is raised considerably above that of the lattice. The amplitude of the field is then reduced to a small value, and observations are made of the exponential growth of signal strength as the nuclear temperature falls to that of the lattice.

4. EXPERIMENTS ON LIQUID AND SOLID ORTHO-HYDROGEN

Measurements were made at a frequency of 5.4 Mc./sec. on the proton resonance in liquid and solid hydrogen. It is probable that the *ortho* : *para* ratio corresponded approximately to the room-temperature equilibrium value of 3 : 1. The values of spin-lattice relaxation time T_1 and line width ΔH which were obtained at different temperatures are given in table 1.

TABLE 1

$T^\circ \text{K}$	20	14.0	11	4.2	1.85	1.3	1.1
T_1 (sec.)	0.02	0.1	0.2	0.3	0.5	2	2
ΔH (gauss)	0.1	0.1	1.0	8	10	50	50

The changes in the width and shape of the line as the temperature is lowered are shown by the oscillograph patterns in figure 3 *a* to *g*. The deflexion in the x direction was proportional to the modulation of H_0 . Two traces are observed, one for each passage through resonance, during the modulation cycle. For the liquid and also for the solid just below the triple point, the width of the line is extremely small ($\simeq 0.1$ gauss), and one observes the transient oscillation effects which are obtained when the time of passage through resonance is short compared with the transverse relaxation time T_2 (Jacobsohn & Wangsness 1948). The width can be calculated from the rate at which the oscillations decay. There is no appreciable change in the resonance pattern on solidification, but, on further cooling, the width increases. At 11.6° K, where the width has increased to about 0.3 gauss, the passage through resonance is sufficiently slow for the transients to be no longer observed. Between 11 and 10° K there is a rapid increase of width to about 8 gauss. Between 10 and 1.7° K the width remains nearly constant. At 1.5° K a further considerable change begins, and, at the lowest temperature which could be attained ($\simeq 1^\circ \text{K}$), there is a very broad resonance with two peaks separated by about 50 gauss.

Discussion of the hydrogen resonance

There are three factors which may influence the width and shape of the resonance line:

- (a) The interaction of the two protons in a molecule with neighbouring molecules.
- (b) The interaction with the internal field due to molecular rotation.
- (c) The mutual interaction of the two protons.

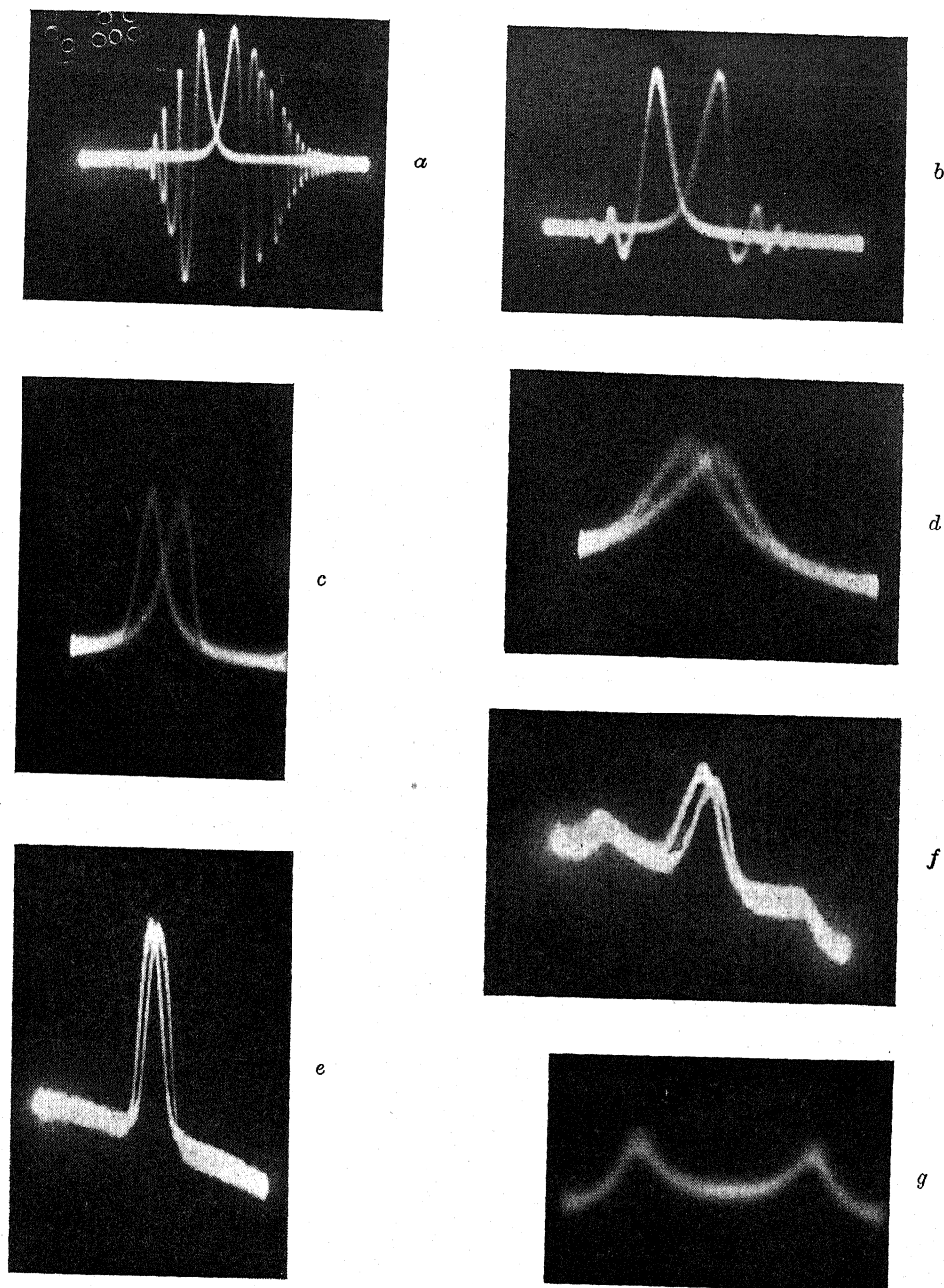


FIGURE 3 (*a*). Liquid hydrogen at 14° K.; width of trace 5 gauss. (*b*) Solid hydrogen at 12.9° K.; width of trace 2.5 gauss. (*c*) Solid hydrogen at 11.6° K.; width of trace 2.5 gauss. (*d*) Solid hydrogen at 11.0° K.; width of trace 2.5 gauss. (*e*) Solid hydrogen at 1.7° K.; width of trace 70 gauss. (*f*) Solid hydrogen at 1.5° K.; width of trace 70 gauss. (*g*) Solid hydrogen at 1° K.; width of trace 85 gauss.

The magnitude of (a) depends on the crystal structures and lattice spacings. Although the structure of solid *ortho*-hydrogen has not been investigated, measurements on *para*-hydrogen at 1.65° K (Keesom, de Smedt & Mooy 1930) have indicated a close-packed hexagonal arrangement with a separation between nearest neighbours of $D = 3.75$ Å. It seems reasonable to assume that the structure and spacing of the 3:1 *ortho-para* mixture is not greatly different from this. In estimating (a) we need take account only of the interaction with the nuclei in neighbouring molecules. Although there may be some contribution due to the molecular rotation moment of the neighbours, this effect is certainly small compared with the nuclear interaction. Using the expression for dipolar broadening in a polycrystalline specimen (Van Vleck 1948) we obtain for the width due to (a)

$$\Delta H = 2.35 \left[\frac{3}{5} \mu^2 \frac{I+1}{I} \sum_K r_{JK}^{-6} \right]^{\frac{1}{2}},$$

where μ is the moment due to the two protons in a molecule, $I = 1$, and r is the intermolecular distance. As the only appreciable contribution to the summation comes from the twelve nearest neighbours this gives

$$\Delta H = 2.35 \left[\frac{36}{5} \frac{I+1}{I} \frac{\mu^2}{D^6} \right]^{\frac{1}{2}} = 8.8 \text{ gauss.}$$

With regard to (b), the field acting on the nuclei due to molecular rotation in the state $J = 1$ for a free molecule is 27 gauss (Kellogg, Rabi, Ramsey & Zacharias 1939). However, the wave functions for a molecule in a crystal may be appreciably different from those for a free molecule, and it is possible that the rotational magnetism may be quenched by the crystalline fields.

A knowledge of the molecular wave functions is also required in order to estimate the magnitude of interaction (c). If the effect of spin-spin interaction is calculated for a free molecule (Kellogg *et al.* 1939), it leads to a shift of the resonance by $\pm 3/5H''$ for the rotational magnetic quantum states $m_J = \pm 1$ and by $\pm 6/5H''$ for $m_J = 0$, where $H'' = \mu_p/r_0^3 = 34$ gauss (r_0 being the internuclear distance and μ_p the proton moment).

If the two protons in a molecule were in fixed positions with the line joining them at an angle θ to H_0 , the shift would be $\pm 3/2H'' (3 \cos^2 \theta - 1)$ (Pake 1948). If we take into consideration that, for a polycrystalline specimen, θ can have all values from 0 to π , then, as shown by Pake, this type of interaction gives a resonance curve which exhibits two peaks separated by $3H''$.

Although, in the case of solid hydrogen, the protons are not in fixed positions, it is probable that, at the lowest temperatures, there is some preferential direction for the molecular axis, determined by the forces exerted by neighbouring molecules, and this will give rise to a resonance curve with two peaks, but it is not possible to calculate the separation.

Turning now to a consideration of the experimental results, the fact that the resonance, just below the triple point, has a width of only about 0.1 gauss, means that all three interactions (a), (b) and (c) average to nearly zero. The fact that both (b) and (c) are suppressed indicates that the direction of the molecular axis is changing

rapidly through all possible positions, and, as the width is also considerably less than that calculated for (a), it must be assumed that there is also considerable freedom of relative motion of the molecules (Bloembergen *et al.* 1948). The existence of this relative motion is, indeed, not surprising, since, on account of the high zero-point energy, the solid has a very open structure, the mean separation between molecules being greater than that corresponding to the minimum of the interaction potential, so that only a small activation energy is required by a molecule in order to be able to move from one lattice position to another. It appears, however, that, at 10° K, this relative motion can no longer take place, as, at this temperature, the full width due to (a) is exhibited.

The effects of (b) and (c) first become apparent at about 1.6° K. It seems probable that the shape of the resonance at the lowest temperature can be attributed to (c). The separation of the two peaks is 50 gauss, or about $1.5H''$.

Some information about the state of order in solid hydrogen can be obtained from specific heat measurements (Mendelssohn, Ruhemann & Simon 1930, 1931; Simon 1930). Experiments have been made on various *ortho-para* mixtures at temperatures down to 2.7° K. It has been found that the specific heat of *ortho*-hydrogen begins to depart from a Debye curve at 11° K, and the anomaly has a maximum at some temperature below 2.7° K.* For a mixture containing 50 % of *ortho*-hydrogen there would be an entropy change of $\frac{1}{2}R \log 3 = 1.1$ cal./mol., due to the ordering of molecular orientation, and $-R \log 0.5 = 1.4$ cal./mol. for the establishment of order in the *ortho-para* arrangement. Between 11 and 2.7° K the total entropy change associated with the anomaly is only about 0.5 cal./mol., so that neither of these processes is completed. From our results it seems probable that there is no freedom of migration of the molecules at temperatures below 10° K, so that the *ortho-para* disorder is apparently 'frozen in'. It seems, furthermore, that the main ordering of the molecular orientation takes place at a temperature in the neighbourhood of 1° K.

5. EXPERIMENTS ON DEUTERIUM

Using a mixture of deuterium with about 10 % hydrogen, we have made observations on the deuteron resonance at temperatures down to 1.1° K. At a resonance frequency of 2.1 Mc./sec. the relaxation time in the liquid was found to be about 1 sec. and the line width was less than 0.1 gauss. In the solid, at a temperature of 1.1° K, the relaxation time was approximately 5 sec. and the line width was about 2 gauss. Thus the line widths are appreciably less and the relaxation times appreciably longer than for the protons in hydrogen at similar temperatures.

6. DISCUSSION OF THE DEUTERON RESONANCE

The ratio of *ortho* to *para* molecules may be assumed to be given by the room-temperature equilibrium value of 2 : 1. At the temperature of the measurements the

* 2.7° K is the lowest temperature at which measurements were made. Further experiments on the specific heat of hydrogen are now in progress.

ortho molecules would be all in the state $J = 0$ and the *para* molecules all in the state $J = 1$. The relative abundance of the different types of molecule would be as follows:

	J	I	abundance (%)
<i>ortho</i>	0	0	11
<i>ortho</i>	0	2	55
<i>para</i>	1	1	33

In considering the ratio of the absorption to be expected from molecules with $I = 2$ to that from molecules with $I = 1$, we have a factor of $5/3$ for the relative abundance and, since the absorption depends on $I(I + 1)$, there is another factor of 3, making altogether a ratio of 5 : 1.

There is no doubt that the resonance we observed was due to the molecules with $I = 2$. Since, for these molecules, $J = 0$, there is no contribution to the width from interactions of the types (b) or (c) considered for the case of *ortho*-hydrogen.

Calculation of the effect due to (a) leads to a width of 2.4 gauss in solid deuterium, which is in reasonable agreement with the experimental value. As would be expected, the electric fields due to near neighbours are too small to produce any appreciable broadening through interaction with the nuclear quadrupole moment.

For molecules with $J = 1$, a broadening would be expected due to (b) and (c) and also because of the interaction of the quadrupole moment with the internal electric field. This latter effect would be the major influence at very low temperatures and would be expected to lead to a line width of the order of 100 gauss. Our experimental arrangement was not sufficiently sensitive to detect the presence of such a broad resonance.

It is to be expected that the relaxation time for the nuclei of *ortho*-deuterium would be longer than in the case of *ortho*-hydrogen, since the relaxation process depends on the interaction of the nuclear moment with magnetic field fluctuations due to the thermal motions of the molecules. Not only is the moment of the deuteron much less than that of the proton, but also, for molecules with $J = 0$, there is no contribution to the field fluctuations due to changes of molecular orientation.

7. NUCLEAR RELAXATION IN IONIC CRYSTALS

Preliminary observations on a number of ionic crystals, in the form of samples compressed from powdered salts, yielded values of the order of a few seconds for the nuclear spin-lattice relaxation times at temperatures in the liquid helium region. In order to investigate the mechanism of the relaxation process, a study was made of the F^{19} nuclei in single crystals of fluorite, at different temperatures and for different values of the magnetic field H_0 . The results of these experiments were as follows:

(a) Colourless fluorite

In table 2 are given the values of T_1 for a colourless transparent crystal of natural fluorite, measured at a frequency of 5.3 Mc./sec. The value of H_0 was 1340 gauss.

(b) Irradiated fluorite

In view of a suggestion by H. Fröhlich that the nuclear spin-lattice relaxation process in ionic crystals might be largely determined by the presence of paramagnetic

impurities or F -centres, the crystal used in experiment (a) was irradiated with X-rays until a blue coloration was produced. It was necessary to carry out the irradiation at a low temperature (90°K) so as to reduce the recombination rate of the F -centres. The values of T_1 for the irradiated crystal, measured again at a frequency of 5.3 Mc./sec. , are shown in table 3. The presence of the F -centres had reduced the relaxation times to approximately one-half of their previous values. In order to investigate the dependence of T_1 on H_0 , measurements were made on the same irradiated crystal under the same conditions but with $\nu_0 = 2.05 \text{ Mc./sec.}$, $H_0 = 503 \text{ gauss}$. The results obtained are given in table 4. It is interesting to note that, if the spin-lattice relaxation were due to the mechanism considered by Waller, T_1 would be proportional to $1/H_0^2$, whereas, in fact, the experiment shows that T_1 increases with increase of H_0 .

TABLE 2

$T^\circ \text{K}$	$T_1 \text{ sec.}$
20	17
4.2	67
1.2	280

TABLE 3

$T^\circ \text{K}$	$T_1 \text{ sec.}$
20	5
4.5	40
1.2	120

TABLE 4

$T^\circ \text{K}$	$T_1 \text{ sec.}$
4.5	20
1.2	60

(c) *Naturally coloured fluorite*

In order to investigate the effect of naturally occurring impurities, observations were made at a frequency of 5.3 Mc./sec. on a crystal of naturally coloured fluorite. The colouring of natural fluorite is commonly due to impurities of rare earths and of the iron group. The values of T_1 in this crystal were found to be very much shorter than for the colourless crystal, as shown in table 5. The resonance width ($\approx 8 \text{ gauss}$) in this specimen was essentially the same as in the colourless crystal.

TABLE 5

$T^\circ \text{K}$	$T_1 \text{ sec.}$
20	0.1
4.2	0.46
1.2	1.0

Using a different crystal of naturally coloured fluorite, measurements of the relaxation time at temperatures down to 0.26°K have been made at a resonance frequency of 1.34 Mc./sec. The values are given in table 6.

It has not yet been possible to determine the impurity content of these specimens, but the presence of paramagnetic impurities in both the clear and coloured fluorite has been definitely proved by measurements of Bleaney and Ingram on paramagnetic resonance absorption at microwave frequencies in which a number of quite intense absorption lines were obtained.

TABLE 6

$T^{\circ}\text{K}$	$T_1\text{ sec.}$
4.2	0.37
1.0	1.5
0.26	5

A quantitative study of the effect of paramagnetic impurities on nuclear relaxation in ionic crystals has been made recently by experiments on the proton resonance due to the water of crystallization of mixed crystals containing known quantities of paramagnetic ions (Bloembergen 1948*a*; Darby & Rollin 1949). It has been established that the relaxation time decreases with increasing paramagnetic content and is approximately inversely proportional to the concentration.

8. DISCUSSION OF THE INFLUENCE OF PARAMAGNETIC IMPURITIES ON NUCLEAR RELAXATION IN IONIC CRYSTALS

In the relaxation mechanism considered by Waller, changes of spin orientation are induced by the fluctuating magnetic fields acting on each spin due to the relative translational motion of neighbouring spins. For nuclear spins this effect is exceedingly small because of the small magnitude of the nuclear moment. The transition probability at low temperatures is about 10^{18} times smaller than for the case of electron spins. Even if there is a high concentration of paramagnetic impurity ions, the field fluctuations due to their translational motions are far too small to account for the observed relaxation times.

Much larger field fluctuations will, however, result from changes of electron-spin orientation caused by spin-spin or spin-lattice interaction. If H_v^2 is the mean square magnetic field fluctuation per unit frequency range in the neighbourhood of ν_0 acting on the nuclear moment, then the probability, W , that a nuclear transition will be induced by the field is given by

$$W \simeq \gamma^2 H_v^2. \quad (6)$$

If the field fluctuations are of a random character with a correlation time τ_c , we may, as in the theory of Bloembergen *et al.* (1948), assume a Debye type of spectrum:

$$H_v^2 = H^2 \frac{4\tau_c}{1 + 4\pi^2\nu^2\tau_c^2}, \quad (7)$$

where H^2 is the total mean square fluctuation. For a nucleus at a distance r from a paramagnetic ion with effective magnetic moment μ_e , the value of H^2 is given by

$$H^2 \simeq \mu_e^2/r^6. \quad (8)$$

In considering the effects of interaction of the electron spins, we may identify τ_c with the spin-spin relaxation time ρ' . The value of ρ' depends on the separation

between the paramagnetic ions, and if the concentration of such ions in the sample is c , we have for reasonable lattice spacings

$$\rho' \simeq 10^{-9}/c \text{ sec.} \quad (9)$$

The expression
$$H_\nu^2 = H^2 \frac{4\rho'}{1 + 4\pi^2\nu^2\rho'^2}, \quad (10)$$

determines the rate at which quanta of energy $h\nu$ can be exchanged with the paramagnetic ions due to the mutual interaction of these ions. However, the theory of paramagnetic spin absorption (Gorter 1947) indicates that the absorption is decreased in the presence of a steady transverse magnetic field H_0 , probably in the ratio

$$\frac{Hi^2}{Hi^2 + H_0^2}, \quad (11)$$

where Hi^2 is the mean square field acting on an ion due to neighbouring ions. It seems plausible to suppose that the absorption of quanta from nuclear transitions is also reduced by this factor. Consequently, from (6), (10) and (11), we obtain for the nuclear transition probability in the presence of the steady field H_0

$$W \simeq \gamma^2 H^2 \frac{4\rho'}{1 + 4\pi^2\nu^2\rho'^2} \frac{Hi^2}{Hi^2 + H_0^2}. \quad (12)$$

If the concentration of paramagnetic ions is c , then

$$Hi^2 \simeq \frac{\mu_e^2}{D^6} c^2, \quad (13)$$

where D is the lattice spacing. The mean square local field at a nucleus distant n lattice spacings from a paramagnetic ion is

$$H^2 \simeq \frac{\mu_e^2}{n^6 D^6}, \quad (14)$$

and so we obtain for the transition probability of such a nucleus, using (9), (12), (13) and (14),

$$W \simeq \gamma^2 \frac{\mu_e^4}{n^6 D^6} c^3 \frac{4 \times 10^{-9}}{(c^2 + 4\pi^2\nu^2 10^{-18})(\mu_e^2 c^2 + D^6 H_0^2)}. \quad (15)$$

If we substitute the numerical values $\gamma = 2 \times 10^4$, $\mu_e^2 = 10^{-39}$, $D = 3 \times 10^{-8}$ we obtain

$$W \simeq \frac{2 \times 10^6}{n^6} c^3 \frac{1}{(c^2 + 4\pi^2\nu^2 10^{-18})(c^2 + 7 \times 10^{-7} H_0^2)}.$$

Hence for a near neighbouring nucleus, for which $n = 1$, we have, with

$$\nu = 5 \times 10^6 \text{ c/sec.}, \quad H_0 = 1500 \text{ gauss and } c = 10^{-3}, \quad W \simeq 2/1.5,$$

so that

$$1/W \simeq T_s \simeq 0.7 \text{ sec.} \quad (16)$$

In order to calculate the effect of electron spin-lattice relaxation on the nuclei, we equate τ_c in equation (7) to the electron spin-lattice relaxation time ρ . Then from

equations (6), (7) and (14) we obtain for the relaxation time of a nucleus adjacent to a paramagnetic ion, the value

$$T_L \simeq \frac{1}{W} \simeq 10^{-15} \frac{1 + 10^{15} \rho^2}{\rho}. \quad (17)$$

For most paramagnetic ions at the temperatures with which we are concerned, $\rho^2 \gg 10^{-15}$, and so we obtain

$$T_L \simeq \rho. \quad (18)$$

For the particular case we have considered, the electron spin-spin and spin-lattice effects would be equal for $\rho \simeq 0.7$ sec., a value which would be expected only at temperatures in the neighbourhood of 0.1° K. It follows that, at all temperatures at which experiments have been made, the electron spin-lattice relaxation will be the mechanism responsible for inducing nuclear relaxation.

So far we have considered only the relaxation time of the nuclei adjacent to paramagnetic ions. For a nucleus distant n lattice spacings from an ion, the relaxation time due to direct interaction with the ion would be n^6 times longer, and the mean relaxation time of all the nuclei would be of the order of that for the most distant ones. This leads to

$$T_{1(\text{direct})} \simeq T_L / c^2. \quad (19)$$

However, these distant nuclei may be able to attain equilibrium more readily by conduction of energy through the nuclear spin system. The nuclear spin-spin interchange time is of the order of 10^{-4} sec., and the average time for quanta to diffuse a distance of n lattice spacings will be about $10^{-4} n^2$. For the nuclei farthest from an ion, $n^2 \simeq c^{-2} / 2.5$, so that the time for diffusion of quanta to the nuclei adjacent to an ion is

$$T_D \simeq 4 \times 10^{-5} c^{-2} \text{ sec.} \quad (20)$$

Hence, for $c = 10^{-3}$, we have $T_D \simeq 4 \times 10^{-3}$ sec., which will usually be shorter than the equilibrium time due to direct interaction given in equation (19).

Since each ion is surrounded by $1/c$ nuclei, the average relaxation time, T_1 , will be the time for exchange of $1/c$ quanta by the nuclei adjacent to the ion. Hence

$$T_1 \simeq \frac{T_L}{c} \frac{1}{z}, \quad (21)$$

where z is the number of near-neighbouring nuclei.

It is possible, however, that the equilibrium rate may be limited by the rate of interchange of quanta between nuclei with $n = 1$ and $n = 2$. In general, an ion will have about ten near neighbouring nuclei, each of which can exchange quanta at a rate of 10^4 per sec., so that the maximum rate at which quanta can be transmitted to the ion is of the order of 10^5 per sec. Hence, with the assumptions we have made, the mean relaxation time cannot be very much shorter than about $10^{-5}/c$ sec. This provides the hold-up, in the example we have chosen, for the case of a specimen at room temperature when $\rho \simeq 10^{-7}$ sec. In this case, therefore, we will have

$$T_1 \simeq 10^{-2} \text{ sec.}$$

On the other hand, at liquid helium temperatures where we may take $\rho \simeq 10^{-2}$ sec., the nuclear spin-lattice relaxation time will be determined by equation (21), and we have

$$T_1 \simeq 1 \text{ sec.}$$

These values are in fairly good agreement with our experimental results.

From equation (21) and succeeding remarks, it will be seen that the nuclear spin-lattice relaxation time is governed by the value of ρ , provided $\rho > 10^{-4}$ sec. For many paramagnetic substances ρ has a value of 10^{-4} sec. at a temperature in the region between 20 and 80° K. At higher temperatures, our theory would suggest that the nuclear spin-lattice relaxation time would be independent of temperature. At lower temperatures, the dependence of T_1 on temperature would be the same as the dependence of ρ .* The dependence on magnetic field is determined by the Debye expression (7), which leads in this case to

$$\frac{1}{T_1} \simeq \frac{H^2}{H_0^2} \frac{1}{\rho},$$

so that the relaxation time depends on H_0 both on account of the field dependence of ρ and because the resonant frequency is proportional to H_0 . Thus at low temperatures we would expect T_1 to increase with decreasing temperature and with increase of the steady magnetic field strength. These deductions are in qualitative agreement with our observations.

9. NUCLEAR RELAXATION IN METALS

(a) Aluminium

The relaxation time of Al^{27} has been measured at a frequency of 2.45 Mc./sec., using a spiral of insulated aluminium foil. The thickness of the foil was of the same order as the penetration depth of the radio-frequency field. The line width was approximately 10 gauss. The values obtained for T_1 are given in table 7.

TABLE 7

$T^\circ \text{ K}$	$T_1 \text{ sec.}$
20	~ 0.2
4.2	~ 1
1.2	~ 1

(b) Copper

Resonance signals due to the isotopes Cu^{63} and Cu^{65} have been observed in a spiral of insulated copper foil, using a frequency of 2.48 Mc./sec. The resonance width for each of the isotopes was approximately 8 gauss. At a temperature of 1.2° K, the value of T_1 for the most abundant isotope, Cu^{63} , was found to be approximately 3 sec. Further experiments at temperatures in the range from 1.1 to 0.27° K have been made on a sample of precipitated copper. The specimen was mixed with alcohol to

* For very low temperatures and low field strength the relaxation process may be determined by electron spin-spin interaction [equation (15)]. The value of T_1 would then be expected to be independent of temperature over a certain range.

provide insulation between the particles. The resonance signal was too weak to permit reliable measurements of the relaxation time to be made, but there was no striking change on cooling from 1.1 to 0.27° K.

The measured values of the relaxation times in metallic aluminium and copper are in approximate agreement with the theory of Heitler & Teller (1936), in which the spin-lattice coupling is assumed to take place via the conduction electrons. However, there are not yet sufficient data to test the theoretical proportionality between $1/T_1$ and T° K.

10. NUCLEAR DEMAGNETIZATION

By taking advantage of the long relaxation time in colourless fluorite it was possible to carry out an experiment on nuclear demagnetization. The crystal was maintained in a field of 4000 gauss for 12 min. at a temperature of 1.2° K, after which the field was reduced rapidly to the resonance value of approximately 500 gauss. The temperature of the nuclear-spin system, measured by the strength of the resonance signal, was found to drop to 0.17° K (the value of H_0/T being approximately constant), and subsequently to rise to the lattice temperature with a relaxation time of 60 sec., in agreement with the previously measured relaxation time.

11. NUCLEAR PARAMAGNETISM AND VERY LOW TEMPERATURES

It is interesting to consider the results on relaxation times in relation to the possible applications of nuclear paramagnetism at very low temperatures.

Owing to the very weak interactions of the nuclei in a substance, the nuclear susceptibility would be expected to obey Curie's law down to temperatures of the order of 10^{-5} ° K, so that measurement of the nuclear susceptibility would provide a convenient way of determining absolute temperatures in this region. For the success of this method it is, of course, essential for the nuclear spin-lattice relaxation time to be reasonably short. From the experimental evidence and also from theoretical considerations it seems unlikely that the use of non-metallic substances, in which the nuclear relaxation is brought about by virtue of electron-spin transitions, will be satisfactory at very low temperatures and in very strong magnetic fields. For metals the situation is more promising, but even in metals we must expect relaxation times at least of the order of several hours at temperatures of 10^{-5} ° K.

By application of a magnetic field of the order of 10^5 gauss at a temperature of 10^{-2} ° K it should be possible to obtain an appreciable degree of nuclear alinement (Kurti & Simon 1935; Simon 1939), and so to investigate the spin dependence of nuclear reactions and the angular distribution of the particles from nuclear disintegrations (Spiers 1948; Rose 1949). By demagnetization from such conditions, extremely low temperatures should be attainable. For the above projects, short relaxation times are again an essential requirement. With values of H_0/T greater than about 10^6 gauss/° K, it would probably not be practicable to rely on the catalytic effect of paramagnetic ions for inducing nuclear relaxation. For metals, for which theory (Heitler & Teller 1936) predicts that the relaxation time varies as

$1/T$ and is independent of H_0 ,* the situation is more hopeful, but the relaxation time would still be very long at the lowest temperatures.

Recently (Gorter 1948; Rose 1949) it has been suggested that it should be possible to achieve alinement of the nuclei of paramagnetic ions by making use of the very large internal field acting on the nucleus due to the electronic moment. By using a mixed salt in which about 10 % of the paramagnetic ions have a nuclear moment, and demagnetizing from an initial temperature of 1° K and a field of 3×10^4 gauss to a temperature of 10^{-20} K and 300 gauss, a considerable degree of nuclear alinement should be obtained. Calculation suggests that in this case the nuclear relaxation time will be quite short, since the value of H_0/T involved is not excessive.

In conclusion, we would like to express our thanks to Lord Cherwell for extending to us the facilities of the Clarendon Laboratory, and to Professor F. E. Simon for his continued encouragement and support of this work. We are greatly indebted to Professor M. H. L. Pryce and Mr K. W. H. Stevens for a number of very helpful discussions, and to Mr J. F. Darby for his able assistance in some of the measurements. One of us (J. H.) would like to make grateful acknowledgement to the Department of Scientific and Industrial Research for a research grant.

REFERENCES

- Bloch, F. 1946 *Phys. Rev.* **70**, 460.
 Bloch, F., Hansen, W. W. & Packard, M. E. 1946 *Phys. Rev.* **70**, 474.
 Bloembergen, N. 1948 *Nuclear magnetic relaxation*. The Hague: Martinus Nijhoff.
 Bloembergen, N. 1948a *Proc. Phys. Soc.* **61**, 575.
 Bloembergen, N., Purcell, E. M. & Pound, R. V. 1948 *Phys. Rev.* **73**, 679.
 Darby, J. F. & Rollin, B. V. 1949 *Nature*, **164**, 66.
 Gorter, C. J. 1934 See Debye, P., *Phys. Z.* **35**, 923.
 Gorter, C. J. 1947 *Paramagnetic relaxation*. Amsterdam: Elsevier.
 Gorter, C. J. 1948 *Physica*, **14**, 504.
 Heitler, W. & Teller, E. 1936 *Proc. Roy. Soc. A*, **149**, 152.
 Jacobsohn, B. A. & Wangsness, R. K. 1948 *Phys. Rev.* **73**, 942.
 Keesom, W. H., de Smedt, J. & Mooy, H. H. 1936 *Proc. Acad. Sci. Amst.* **33**, 814.
 Kellogg, J. M. B., Rabi, I. I., Ramsey, N. F. & Zacharias, J. R. 1939 *Phys. Rev.* **56**, 728.
 Kurti, N. & Simon, F. 1935 *Proc. Roy. Soc. A*, **155**, 629.
 Mendelssohn, K., Ruhemann, M. & Simon, F. 1930 *Naturwissenschaften*, **18**, 34.
 Mendelssohn, K., Ruhemann, M. & Simon, F. 1931 *Z. phys. Chem. B*, **15**, 121.
 Pake, G. E. 1948 *J. Chem. Phys.* **16**, 327.
 Purcell, E. M., Torrey, H. C. & Pound, R. V. 1946 *Phys. Rev.* **69**, 37.
 Rollin, B. V. 1946 *Nature*, **159**, 201.
 Rollin, B. V. & Hatton, J. 1947 *Nature*, **159**, 201.
 Rollin, B. V. & Hatton, J. 1948 *Phys. Rev.* **74**, 346.
 Rollin, B. V., Hatton, J., Cooke, A. H. & Benzie, R. J. 1947 *Nature*, **160**, 436.
 Rose, M. E. 1949 *Phys. Rev.* **75**, 213.
 Simon, F. 1930 *Ergebn. Exakt. Naturw.* **9**, 269.
 Simon, F. 1939 *Conference on magnetism, Strasbourg*, **3**, 1.
 Spiers, J. A. 1948 *Nature*, **161**, 807.
 Van Vleck, J. H. 1948 *Phys. Rev.* **74**, 1168.
 Waller, I. 1932 *Z. Phys.* **79**, 370.

* [Note added in proof.] In recent experiments we have found that for aluminium and copper $1/T_1$ is proportional to T° K, in the temperature range from 20° K to 1° K. The value of T_1 appears to increase with increase of H_0 .

The nature of turbulent motion at large wave-numbers

BY G. K. BATCHELOR, *Trinity College, University of Cambridge*
AND A. A. TOWNSEND, *Emmanuel College, University of Cambridge*

(Communicated by H. Jeffreys, F.R.S.—Received 4 May 1949)

This paper presents some measurements which describe the distribution of turbulent energy over Fourier components of large wave-number. According to Kolmogoroff's theory these components of the motion are in statistical equilibrium. Over part of the wave-number range the form of the spectrum is predicted by a dimensional argument, but at higher wave-numbers the spectrum depends on the manner in which energy is transferred across the spectrum. Several postulates about this transfer effect have been made. Only that made by Heisenberg leads to a spectrum function which can be consistent with the measurements. Consistency with Heisenberg's spectrum is obtained by postulating that there is an upper limit to the range of wave-numbers containing energy. This limit evidently corresponds to a critical Reynolds number below which no energy transfer occurs and the motion is 'laminar'.

Measurements describing the probability distribution of $\partial u/\partial x$, $\partial^2 u/\partial x^2$ and $\partial^3 u/\partial x^3$ are also described. These, and oscillograms of the velocity derivatives, show that the energy associated with large wave-numbers is very unevenly distributed in space. There appear to be isolated regions in which the large wave-numbers are 'activated', separated by regions of comparative quiescence. This spatial inhomogeneity becomes more marked with increase in the order of the velocity derivative, i.e. with increase in the wave-number. It is suggested that the spatial inhomogeneity is produced early in the history of the turbulence by an intrinsic instability, in the way that a vortex sheet quickly rolls up into a number of strong discrete vortices. Thereafter the inhomogeneity is maintained by the action of the energy transfer.

1. INTRODUCTION

Kolmogoroff's theory of local similarity of turbulent motion (Kolmogoroff 1941; Batchelor 1947) has received strong experimental support (Townsend 1948*a*, *b*) and may now be regarded as an accurate description of one part of the motion. This theory states that the small-scale structure of turbulent motion has invariant statistical properties, apart from changes in the length and velocity scales, under different conditions. Whatever the nature of the large-scale structure of the turbulence, the small-scale structure is postulated to be statistically isotropic and approximately homogeneous in both space and time. The meaning of the latter property is that the statistical characteristics of the small-scale motion are approximately unchanged by translations in space and time over intervals which are large compared with lengths and periods typical of the small-scale structure. From a consideration of the way in which energy is transmitted from large-scale fluctuations of the motion, it is postulated that the statistical properties of the small-scale motion are determined by the two parameters ν (kinematic viscosity of the fluid) and ϵ (energy dissipation per unit mass of fluid).

The theory is best formulated analytically in terms of the energy-spectrum function, since the physical significance of this quantity can easily be understood. In particular, physical understanding is used to justify the assumption that this spectrum function at wave-number k is determined largely by that part of the turbulent motion which is associated with the length scale $2\pi/k$. (An equivalent

assumption has previously been made about the correlation function (Kolmogoroff 1941), and while there are circumstances in which it is probably accurate for both functions, the authors incline to the belief that it is more accurate for the spectrum function.) If

$$R_i^j(\mathbf{r}) = \overline{u_i(\mathbf{x}) u_j(\mathbf{x} + \mathbf{r})}, \quad (1.1)$$

where $u_i(\mathbf{x})$ is the i -component of velocity at a point specified by the vector $\mathbf{x}$, the general three-dimensional spectrum tensor is defined (Batchelor 1949) as

$$\Gamma_i^j(\mathbf{k}) = \frac{1}{8\pi^3} \int \int \int_{-\infty}^{\infty} e^{-i(\mathbf{k} \cdot \mathbf{r})} R_i^j(\mathbf{r}) d\tau(\mathbf{r}), \quad (1.2)$$

where $d\tau(\mathbf{r})$ is an element of volume at a point specified by $\mathbf{r}$, and $\mathbf{k}$ is a vector wave-number. The corresponding transform relation

$$R_i^j(\mathbf{r}) = \int \int \int_{-\infty}^{\infty} e^{i(\mathbf{k} \cdot \mathbf{r})} \Gamma_i^j(\mathbf{k}) d\tau(\mathbf{k}) \quad (1.3)$$

shows that $\frac{1}{2}[\Gamma_i^j(\mathbf{k})]_{i=j}$ is the density in wave-number space of contributions to the energy of the fluctuation in the i -direction, and that $\frac{1}{2}\Gamma_i^i(\mathbf{k})$ (repeated indices summed) is the density of contributions to the energy of all three velocity components.

The contribution to the energy from wave-numbers whose moduli lie between k and $k + dk$ is thus $E(k) dk$, where

$$\begin{aligned} E(k) &= \frac{1}{2} \int_0^\pi \int_0^{2\pi} \Gamma_i^i(\mathbf{k}) k^2 \sin \theta' d\theta' d\phi' \\ &= \frac{1}{4\pi^2} \int_0^\infty \int_0^\pi \int_0^{2\pi} R_i^i(\mathbf{r}) kr \sin kr \sin \theta'' dr d\theta'' d\phi'', \end{aligned} \quad (1.4)$$

where (k, θ', ϕ') and (r, θ'', ϕ'') are spherical polar co-ordinates of the vectors $\mathbf{k}$ and $\mathbf{r}$. When the turbulence is statistically isotropic, $R_i^i(\mathbf{r})$ can be expressed in terms of a single scalar function $\overline{u^2}f(r)$:

$$R_i^i(\mathbf{r}) = \overline{u^2} \left[3f(r) + r \frac{\partial f(r)}{\partial r} \right], \quad (1.5)$$

so that

$$\begin{aligned} E(k) &= \frac{\overline{u^2}}{\pi} \int_0^\infty kr \sin kr \left(3f + r \frac{\partial f}{\partial r} \right) dr \\ &= \frac{\overline{u^2}}{\pi} \int_0^\infty f(r) k^2 r^2 \left(\frac{\sin kr}{kr} - \cos kr \right) dr. \end{aligned} \quad (1.6)$$

The transformed relation corresponding to (1.6) is

$$\overline{u^2}f(r) = 2 \int_0^\infty E(k) k^{-2} r^{-2} \left(\frac{\sin kr}{kr} - \cos kr \right) dk. \quad (1.7)$$

These equations relate the quantity under discussion, $E(k)$, to the quantity $f(r)$, which is sometimes easier to measure.

Now consider the implications of Kolmogoroff's theory for the spectrum function $E(k)$. Inasmuch as $E(k)$ is a quantity determined by that part of the turbulence which is associated with the length scale $2\pi/k$ (as assumed above), it will come within the scope of the theory when k is large enough and must then be a similar function under all conditions. Changes in the length and velocity scales are admitted, these being determined by the values of the controlling parameters ν and ϵ . Hence when k is large enough,

$$E(k) = \epsilon^{\frac{1}{3}} \nu^{\frac{1}{3}} F(k\nu^{\frac{1}{3}}\epsilon^{-\frac{1}{3}}), \quad (1.8)$$

where $F(\xi)$ is a *universal* function of its argument ξ . The range of values of k for which Kolmogoroff's theory (and, more specifically, the prediction (1.8)) is accurate is termed the *equilibrium range* of wave-numbers, because these wave-numbers can be regarded as degrees of freedom which are in mutual equilibrium under the action of inertial and viscous forces. Likewise the function F describes the *equilibrium spectrum*. The lower bound ($k = k_1$, say) of the equilibrium range will depend on the type of turbulence under consideration. In the common case of decaying isotropic turbulence k_1 must be such that most of the energy of the turbulence lies in wave-numbers less than k_1 , since otherwise the equilibrium could not be approximately steady. In the other direction the equilibrium range is usually assumed (tacitly) to extend to infinite wave-numbers; this is a point about which more will be said later.

In view of the comparative simplicity of conditions in the equilibrium range, it is natural to concentrate research on this aspect of the problem of turbulence. Moreover, the universality of the equilibrium spectrum gives great importance to any results which can be obtained. The object of this paper is to present some measurements relating to the equilibrium spectrum and to consider their implications.

2. THE FORM OF THE EQUILIBRIUM SPECTRUM

In his original paper, Kolmogoroff (1941) postulated that when the Reynolds number of the turbulence is high enough there is a subrange of the equilibrium range in which conditions are not affected by viscosity. The physical idea behind this postulate is that the motion associated with a small range of wave-numbers has its own Reynolds number and at smaller values of k for which this characteristic Reynolds number is very large the motion is dominated by inertia forces. Within this subrange of k -values, the universal function F must have a form such that ν disappears from the expression (1.8) for $E(k)$. Hence when $k_1 < k \leq k_R$, where k_R depends on (and increases with) Reynolds number of the turbulence, we have

$$E(k) = A\epsilon^{\frac{2}{3}}k^{-\frac{5}{3}}, \quad (2.1)$$

where A is an absolute constant. k_R is a wave-number in the range in which the effects of viscous forces on the spectrum first become comparable with the effects of inertia forces. Since this range is entirely determined by ν and ϵ , k_R is proportional to $(\epsilon/\nu^3)^{\frac{1}{3}}$; the constant of proportionality is of order unity and k_R may therefore be defined as

$$k_R = \left(\frac{\epsilon}{\nu^3}\right)^{\frac{1}{3}}. \quad (2.2)$$

The form of the double-velocity correlations corresponding to (2.1), and the form of the triple-velocity correlation found from the same postulates about the inertial subrange, have each been found to be in good agreement with measurements made at high Reynolds number (Townsend 1948*a*); and the result (2.1) may be accepted as accurate.

In the range of wave-numbers $k_R < k < \infty$, both inertia and viscous forces are operative, and dimensional analysis is not sufficient to determine the form of the equilibrium spectrum. To deduce the equilibrium spectrum in this range something must be known about the way in which energy is transferred between different wave-numbers. *Ab initio* determinations of the non-linear transfer effect are practically ruled out by the complexity of the problem, and several authors have therefore been led to make intuitive postulates. In the opinion of the present authors Heisenberg's is the only one of these postulates which has any *a priori* likelihood of leading to a good approximation to the spectrum, and even here the analysis is plausible rather than convincing. However, an assessment of the relative merits of the different postulates about the energy transfer is outside the scope of this paper, and here the forms of the equilibrium spectrum put forward by various authors are simply quoted without comment on their derivations.

(a) Obukhoff (1941) finds, as a consequence of an assumption about the transfer of energy between wave-numbers, that the equilibrium spectrum is given by the equation

$$\epsilon = 2\chi \left[\int_0^k k^2 E(k) dk \right]^{\frac{1}{2}} \int_k^\infty E(k) dk + 2\nu \int_0^k k^2 E(k) dk, \quad (2.3)$$

where χ is a universal constant for sufficiently large Reynolds numbers. An explicit solution for $E(k)$ is not given, although, in common with all dimensionally correct arguments, Obukhoff's postulate leads to the inertial solution (2.1) for smaller values of k within the equilibrium range. The spectrum function given by (2.3) has a minimum in the neighbourhood of $k = \sqrt{(\chi)} k_R$, so that it is doubtful whether it ought to be taken literally outside the inertial subrange (where, in any case, no assumption about the transfer term is needed). Obukhoff's interpretation is that the spectrum contains no energy at wave-numbers greater than $\sqrt{(\chi)} k_R$.

(b) Heisenberg (1948*a*) makes a different postulate about the transfer effect which leads to the result

$$E(k) = \left(\frac{8\epsilon}{9k} \right)^{\frac{1}{2}} k^{-\frac{1}{2}} \left[1 + \left(\frac{k}{k_s} \right)^4 \right]^{-\frac{1}{2}} \quad (2.4)$$

within the equilibrium range, where $k_s = \left(\frac{3\kappa^2}{8} \right)^{\frac{1}{2}} k_R$ and κ is an absolute constant.

When $k \gg k_s$ (2.4) reduces to $E(k) = \left(\frac{\kappa\epsilon}{2\nu^2} \right)^{\frac{1}{2}} k^{-7}$. Heisenberg mentions the possibility that the motion is not 'turbulent' at very high wave-numbers, and consequently that (2.4) is not valid for indefinitely high values of k .

(c) Yet another assumption about the energy transfer has been made by Kovasznay (1948), and the resulting expression for $E(k)$ is

$$E(k) = A\epsilon^{\frac{1}{2}} k^{-\frac{1}{2}} \left[1 - \left(\frac{k}{k_0} \right)^{\frac{1}{2}} \right]^2 \quad (k \leq k_0), \quad (2.5)$$

where k_0 is proportional to k_R . The characteristic feature of this spectrum is that there is no energy associated with wave-numbers greater than a critical value proportional to k_R .

It has been found convenient experimentally to obtain information about the spectrum by making measurements of integral moments of $E(k)$ —or, what amounts to the same thing, measurements of derivatives of the correlation function $\overline{u^2 f(r)}$. The relation (1.7) shows that

$$D_{2n} = \overline{u^2} \left[\frac{\partial^{2n} f(r)}{\partial r^{2n}} \right]_{r=0} = \frac{2(-1)^n}{(2n+3)(2n+1)} \int_0^\infty k^{2n} E(k) dk. \quad (2.6)$$

The dependence of these integral moments on the form of the equilibrium spectrum will vary with n . It is clear that measurements of $D_0 (= \overline{u^2})$ provide no information about the equilibrium spectrum, since by definition the equilibrium spectrum contains a negligible proportion of the total energy. The next derivative, D_2 , is proportional to the total viscous dissipation of energy. Taylor (1938) has pointed out that when the Reynolds number is sufficiently large the energy of the turbulence and the viscous dissipation in a decaying isotropic field are associated with different, approximately non-overlapping, ranges of wave-numbers. In these circumstances, i.e. when $k_R > k_1$, the quantities $D_2, D_4, \dots$ are determined almost wholly by the form of the equilibrium spectrum and the parameters ϵ and ν . In particular, the dimensionless ratio $D_6 D_4^{-2} D_2$ is then an absolute constant. Equations (1.8) and (2.6) together give

$$D_6 D_4^{-2} D_2 = \frac{35}{27} \frac{\int_0^\infty \xi^6 F(\xi) d\xi \int_0^\infty \xi^2 F(\xi) d\xi}{\left[\int_0^\infty \xi^4 F(\xi) d\xi \right]^2}, \quad (2.7)$$

provided $k_R > k_1$.

Measurements of this quantity are presented later, and in the meantime, the predictions of its value made by the suggested spectrum functions (2.3), (2.4) and (2.5) may be noted. If Obukhoff's spectrum is assumed to terminate at $k = \sqrt{(\chi)} k_R$, the consequent value of $D_6 D_4^{-2} D_2$ is found to be 1.83; this figure is not altered very much by any other choice of the upper terminal wave-number. Heisenberg's equilibrium spectrum (2.4), if taken literally, gives an infinite value of D_6 and of $D_6 D_4^{-2} D_2$. However, as already mentioned, Heisenberg (1948*a*) states that (2.4) may not be valid for indefinitely large values of k because there is likely to be a critical value of the Reynolds number associated with large wave-numbers below which no further transfer of energy occurs.* If this is so, and if we assume the spectrum (2.4) to be cut off sharply at $k = k_2$, the consequent value of $D_6 D_4^{-2} D_2$ increases with k_2 in a manner which needs to be determined numerically. Calculations show, for example, that the value of $D_6 D_4^{-2} D_2$ is about 1.9 when $k_2/k_s = 10^\dagger$ and about 3.8 when $k_2/k_s = 100^\dagger$. Kovasznay's equilibrium spectrum (2.5) leads to $D_6 D_4^{-2} D_2 = 2.79$. In general, any

* In an appendix to this paper, another possible reason, suggested previously by one of the present authors, for the existence of a finite upper limit to the equilibrium range of wave-numbers in decaying turbulence is discussed and is shown not to impose any restriction on the range of applicability of Heisenberg's spectrum.

reasonable shape of spectrum, for which the integral D_6 converges, leads to a value of $D_6 D_4^{-2} D_2$ which lies between 1.4 and 3.* For instance, the simple power law $E \propto k^{-n}$ leads to values 2.33, 1.73 and 1.54 when n is 8, 9 and 10 respectively. If E has the form $k^n e^{-2\alpha k^2}$, then $D_6 D_4^{-2} D_2$ takes the values 2.59, 2.16, 1.95 and 1.67 when n is -1, 0, 1 and 4 respectively. It will be shown later that the experimental value at high Reynolds number is a good deal larger than any of these predictions, except that obtained from a spectrum of Heisenberg's form which is terminated at an appropriately chosen wave-number.

3. EXPERIMENTAL METHOD

The measurements described below were all made in the small wind-tunnel in the Cavendish Laboratory, details of which have been given previously (Batchelor & Townsend 1947). With the exception of the hot-wire amplifier, the hot-wire anemometer and its associated electrical circuits used in the present experiments are the same as those described in that paper. The amplifier (figure 1) has been modified to

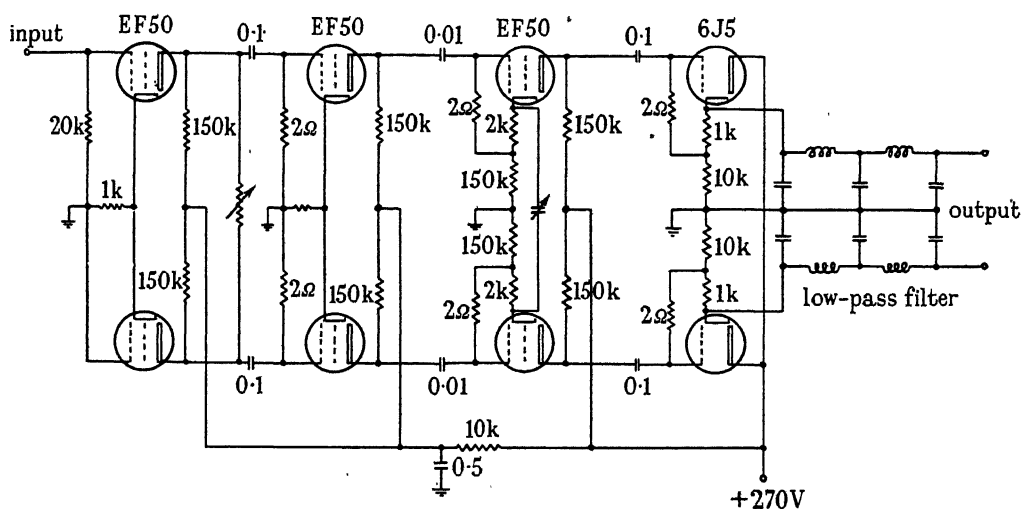


FIGURE 1. Turbulence amplifier with low-pass filter.

($1\omega = 1$ ohm, $1k = 10^3$ ohm, $1\Omega = 10^6$ ohm.)

allow the limitation of the upward frequency response by plugging in low-pass filters of appropriate cut-off frequencies. In this way it has been possible to adjust the frequency range of the amplifier so that loss owing to finite frequency range is negligible, but, at the same time, very little electrical noise of unwanted frequencies is amplified. The effectiveness of this arrangement is such that it is possible to differentiate the hot-wire signal three times before the amplifier noise becomes appreciable. From the amplifier, the hot-wire output is led into a circuit which may be set to give an output proportional to the original signal, or to its first,

* Note added in proof: A spectrum function for which the integral moments D_{2n} are slowly convergent leads to values of $D_6 D_4^{-2} D_2$ outside this range. For instance, $E(k) \propto k^{-\alpha} \log k$ leads to $D_6 D_4^{-2} D_2 = 9.57/\alpha$.

second or third derivative with respect to time, the mean square value then being determined in the usual way with a thermo-junction and millivoltmeter. The differentiation with respect to time is performed by circuits similar to those described by one of the authors (Townsend 1947).

The determination of D_2 , D_4 and D_6 proceeds from their expression as mean squares of derivatives of the velocity:

$$D_{2n} = \overline{u^2 \left[\frac{\partial^{2n} f(r)}{\partial r^{2n}} \right]_{r=0}} = (-1)^n \overline{\left(\frac{\partial^n u}{\partial x^n} \right)^2}. \quad (3.1)$$

The experimental measurement of these quantities depends on the approximation that spatial differentiation with respect to x (i.e. in the downstream direction) may be replaced by differentiation with respect to time at a point fixed with respect to the wind tunnel. Symbolically, this means

$$\frac{\partial}{\partial x} = -U \frac{\partial}{\partial \tau},$$

where U is the mean velocity and τ is the current time for the hot-wire record. This assumption is clearly valid in the limit when U is large compared with any turbulent velocity fluctuation. The justification of its use in the present and previous experimental work rests mainly in the consistency of the results, and the excellent agreement of observed values of $\overline{\left(\frac{\partial u}{\partial x} \right)^2}$ and those calculated from the energy equation, viz.

$$\frac{d\overline{u^2}}{dt} = -10\nu \overline{\left(\frac{\partial u}{\partial x} \right)^2}.$$

Assuming the equivalence of space and time differentiation, x may be replaced by τ in the expression for D_{2n} , and the quantity to be measured is now $\overline{\left(\frac{\partial^n u}{\partial \tau^n} \right)^2}$ for $n = 1, 2$ and 3 .

Since the values of high-order derivatives of the velocity fluctuation are controlled by the spectrum components of very high wave-number, losses must be expected either if the hot wire is too long or if the frequency range of the amplifier is insufficient. In the present experiments, the effect of wire length was investigated by repeating the observations with a longer wire, and the absence of a systematic difference between the two sets of observed values of D_4 and D_6 was interpreted as evidence that hot-wire length is not a serious limitation in these experiments. Similarly, by varying the frequency range of the amplifier, a frequency range sufficient to include all the relevant Fourier components of the turbulence was determined, and used for the final measurements. In this way, the authors satisfied themselves that further extension of the frequency range of the amplifier or further reduction in length of the hot wire would not appreciably change the final results.

Some measurements of the flattening factor of the velocity fluctuation and of the first three derivatives of the velocity fluctuation were also made, using ordinary technique (Townsend 1948*b*).

4. EXPERIMENTAL RESULTS

Most of the measurements were made in isotropic turbulence produced by inserting square-mesh grids (made of circular rods whose diameter was $\frac{3}{16}$ of the periodic spacing M) in a uniform airstream. M took the values 0.64, 1.27, 2.54 and 5.08 cm. and the stream velocity U was either 300 or 620 cm.sec.⁻¹. For any grid Reynolds number $R_M (= UM/\nu)$ it was found that the dimensionless quantities

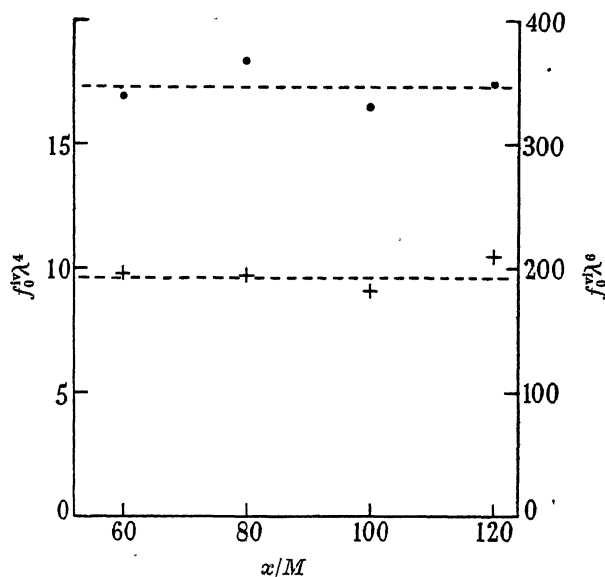


FIGURE 2. Variation of $f_0^{iv} \lambda^4$ (•) and $f_0^{vi} \lambda^6$ (+) with x/M . $R_M = 5620$.

$f_0^{iv} \lambda^4$ and $f_0^{vi} \lambda^6$ were constant during the decay process (figure 2), where λ is the dissipation length parameter defined by $f_0'' = -\lambda^{-2}$. The constancy of $f_0^{iv} \lambda^4$ has been reported in a previous paper (Batchelor & Townsend 1947), and both results are consistent with the postulated theory of the initial period of decay (Heisenberg 1948*b*).

Another, although less accurate, method of determining $f_0^{iv} \lambda^4$ and $f_0^{vi} \lambda^6$ would be to find the polynomials of the second and third degree in $(r/\lambda)^2$ which give the best fit with the measured correlation function. Figure 3 shows a comparison of the measured correlation function $g(r) = f + \frac{1}{2} r \partial f / \partial r$ with the two polynomials

$$1 - \left(\frac{r}{\lambda}\right)^2 + \frac{3}{4!} f_0^{iv} \lambda^4 \left(\frac{r}{\lambda}\right)^4$$

and

$$1 - \left(\frac{r}{\lambda}\right)^2 + \frac{3}{4!} f_0^{iv} \lambda^4 \left(\frac{r}{\lambda}\right)^4 + \frac{4}{6!} f_0^{vi} \lambda^6 \left(\frac{r}{\lambda}\right)^6,$$

using the directly measured values of $f_0^{iv} \lambda^4$ and $f_0^{vi} \lambda^6$. The agreement provides a check on the general validity of the direct measurements of $f_0^{iv} \lambda^4$ and $f_0^{vi} \lambda^6$.

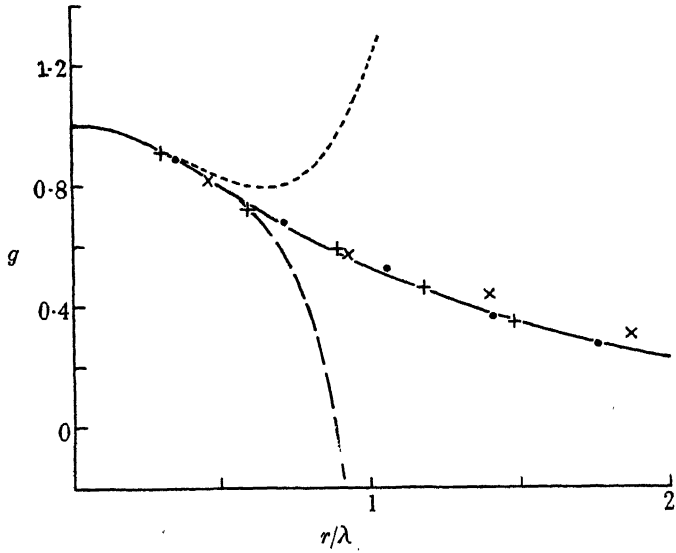


FIGURE 3. Transverse correlation.

$$R_M = 5620, \quad f_0^{iv}\lambda^4 = 9.5, \quad f_0^{vi}\lambda^6 = 340.$$

$$\times, x/M = 40; \quad \bullet, x/M = 60; \quad +, x/M = 80.$$

$$\cdots \cdots \cdots 1 - \frac{r^2}{\lambda^2} + \frac{f_0^{iv}\lambda^4}{8} \frac{r^4}{\lambda^4}.$$

$$- - - - - 1 - \frac{r^2}{\lambda^2} + \frac{f_0^{iv}\lambda^4}{8} \frac{r^4}{\lambda^4} - \frac{f_0^{vi}\lambda^6}{180} \frac{r^6}{\lambda^6}.$$

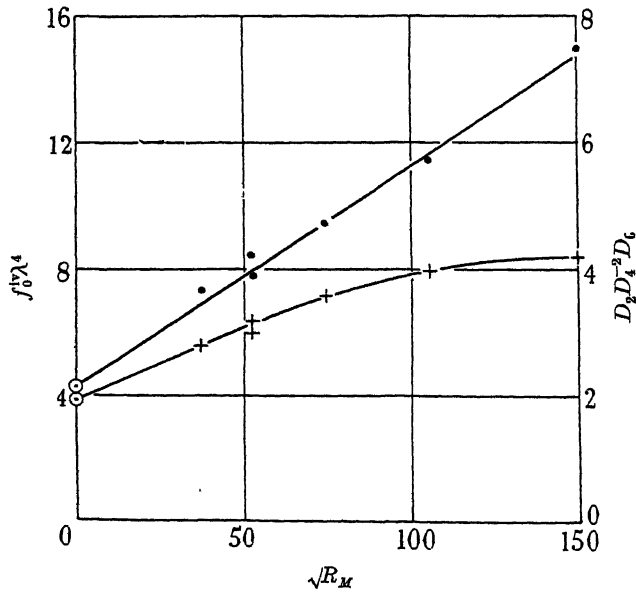


FIGURE 4. Variation of $f_0^{iv}\lambda^4$ and $D_2D_4^{-2}D_6$ with Reynolds number.
 $\bullet, f_0^{iv}\lambda^4$; $+, D_2D_4^{-2}D_6$; $\circ$, theoretical value at low Reynolds number.

However, the quantities $f_0^{iv}\lambda^4$ and $f_0^{vi}\lambda^6$ are both influenced by the total energy of the turbulence, and it is the ratio

$$\frac{f_0^{vi}\lambda^6}{(f_0^{iv}\lambda^4)^2} = D_6 D_4^{-2} D_2$$

which is of interest in this paper. The results just mentioned show that it is constant during decay, and figure 4 shows how it varies with the grid Reynolds number. (The abscissa used in this diagram is $\sqrt{R_M}$, since the variation of $f_0^{iv}\lambda^4$ is then linear, as has already been demonstrated (Batchelor & Townsend 1947).) It was shown earlier that the theory of the equilibrium spectrum requires this quantity to be invariant at sufficiently high Reynolds number, and the experimental results show a tendency in this direction for R_M greater than about 1.5×10^4 . This is about the value of R_M above which we should expect $D_6 D_4^{-2} D_2$ to be invariant, judging from measurements of the correlation function reported in a previous paper (Batchelor & Townsend 1948). It was shown there that the correlation function first begins to approximate to the form $1 - f(r) \propto r^3$ for small values of r , at a Reynolds number R_M between 10^4 and 4×10^4 . This form for $f(r)$ is the counterpart of the expression (2.1) for $E(k)$, both expressions being consequences of the existence of an inertial subrange of the equilibrium range of wave-numbers, and its appearance should mark approximately the Reynolds number at which $D_6 D_4^{-2} D_2$ becomes invariant.

At Reynolds numbers smaller than about 10^4 , $D_6 D_4^{-2} D_2$ is not invariant, and its value is determined by the shape of the spectrum in the quasi-equilibrium range of wave-numbers. Within this range the spectrum shape belongs to a singly infinite family of curves, with the Reynolds number as parameter (Heisenberg 1948*b*). Once again determination of the spectrum requires a knowledge of the inertial transfer of energy across the spectrum. However, in the case of very low Reynolds number, the direct effect of inertia forces becomes negligibly small over the quasi-equilibrium range, and it is possible to make a definite prediction, viz.

$$E(k) = \text{const. } k e^{-2\nu k^2}. \quad (4.1)$$

The corresponding theoretical correlation function has previously been found to agree well with measurements (Batchelor & Townsend 1948). The expression (4.1) will not be correct for very small values of k , but there will be very little consequent error in the calculation of $D_6 D_4^{-2} D_2$. Nor will it be correct at very high wave-numbers where the equilibrium form of the spectrum begins to be valid. The equilibrium range moves to higher wave-numbers as the Reynolds number decreases, so that ultimately the integrals D_2 , D_4 and D_6 will be determined wholly by the quasi-equilibrium form (4.1) (always assuming that the integrals converge—for the spectrum given by equation (2.4) the contribution to D_6 from the equilibrium range is infinite, and remains so at all Reynolds numbers). Thus in the limit of very small Reynolds number (4.1) and (2.6) show that

$$D_6 D_4^{-2} D_2 = \frac{3.5}{1.8} = 1.94,$$

and this prediction is consistent with the trend of the curve in figure 4.

The asymptotic value of $D_6 D_4^{-2} D_2$ at large Reynolds number appears to be at least as large as 4.2. This is well above the value predicted by any of the spectrum curves described earlier, except in the case of Heisenberg's spectrum for which the integral D_6 increases indefinitely with increase of the upper terminal. Calculations as in § 2 show that Heisenberg's spectrum is consistent with the asymptotic value of $D_6 D_4^{-2} D_2$, being equal to 4.2, provided one can assume that it applies for $k < 3.5k_s$, and that there is no energy in larger wave-numbers. Unless some other possible explanation of the high measured value of $D_6 D_4^{-2} D_2$ can be found, the evidence definitely favours Heisenberg's form of the equilibrium spectrum, despite the presence of the theoretically arbitrary cut-off wave-number.

However, it is now very important to know more about the way in which the spectrum terminates. Heisenberg's idea (expressed very briefly in his paper) appears to be that there is a critical value of some Reynolds number characteristic of the very

small eddies, say $R = \frac{\left(\int_k^\infty E(k) dk\right)^{\frac{1}{2}}}{\nu k}$, below which energy transfer is much smaller than his transfer term predicts and the spectrum falls off sharply. Thus the critical value of R which is consistent with the assumption that Heisenberg's spectrum (2.4) is valid for $k < 3.5k_s$ and that $E = 0$ for higher wave-numbers is given by

$$R_c = \frac{1}{\sqrt{6}} \frac{\kappa \epsilon}{\nu} \frac{1}{2\nu^2 (3.5k_s)^4} \\ = \kappa^{-1} \left(\frac{2}{3}\right)^{\frac{1}{2}} (3.5)^{-4} = 0.0045,$$

since the value of κ is in the neighbourhood of 0.8.

The strangely high value of $D_6 D_4^{-2} D_2$ suggested to the authors that other statistical properties of the motion associated with large wave-numbers should be examined.

One of the simplest of these is the flattening factor $\alpha_n = \left(\frac{\partial^n u}{\partial x^n}\right)^4 / \left[\left(\frac{\partial^n u}{\partial x^n}\right)^2\right]^2$ of the velocity derivative $\partial^n u / \partial x^n$. Owing to the presence of two mean values in the denominator of α_n as against one in the numerator, α_n has the property that it is sensitive to any appreciable lack of compactness of the probability distribution of $\partial^n u / \partial x^n$. For instance, to take an extreme case, if $|\partial^n u / \partial x^n|$ had the value C with probability γ and the value 0 with probability $1 - \gamma$, we should have $\alpha_n = \gamma^{-1}$. Generally speaking, the flattening factor is increased by any tendency for the probability distribution to approximate to the occurrence of either one of two values of $|\partial^n u / \partial x^n|$, one of which is large compared with the other.

Measurements of α_n for $n = 0, 1, 2$ and 3 at various Reynolds numbers are shown in figure 5. The flattening factor for u is close to 3.0, which is the figure for a Gaussian probability distribution of u . With higher values of n , α_n seems to vary a little with Reynolds number. α_n ($n > 0$) should be determined wholly by the equilibrium range of wave-numbers at high enough Reynolds numbers and should then be an absolute constant; possibly the range of Reynolds numbers of the experiments is not great enough to show this result properly. The interesting aspect of figure 5 is the rapid increase of α_n with n . The probability distributions of $\partial^2 u / \partial x^2$ and $\partial^3 u / \partial x^3$ must have

very unusual shapes to give values of the flattening factor near 4.9 and 5.9. On the simple assumption that $\partial^n u / \partial x^n$ is zero for a fraction $1 - \gamma$ of the total time and varies with a Gaussian probability distribution during the remainder of the time, flattening factors 3.0, 3.9, 4.9 and 5.9 (for $n = 0, 1, 2$ and 3 respectively) correspond to $\gamma = 1.0, 0.77, 0.61$ and 0.51 respectively. Since high orders of the velocity derivative are associated with very large wave-numbers, the results of figure 5 suggest that there is present an effect which becomes increasingly important as the wave-number is increased.

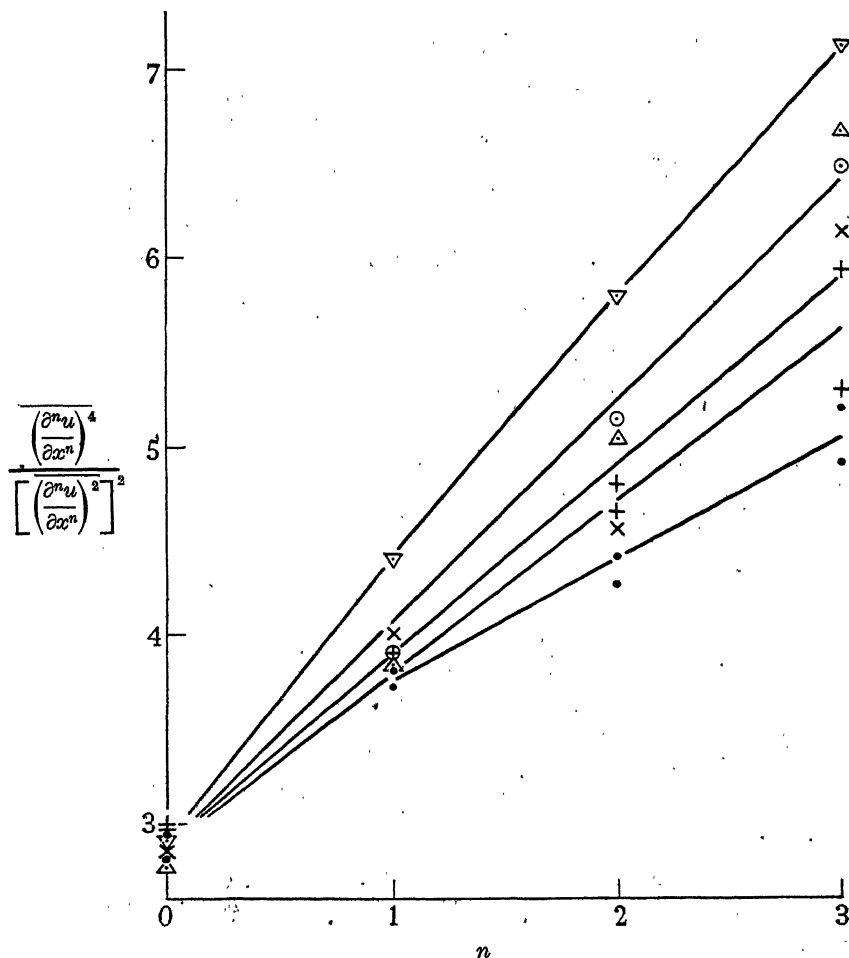


FIGURE 5. Flattening factors of velocity derivatives.

Isotropic turbulence, R_M : •, 2810; +, 5620; x, 11,200; ◯, 22,500.

Cylinder wake, R_d : Δ, 680; ∇, 4,100.

In view of the above-mentioned property of the flattening factor, the inference is that $\partial^n u / \partial x^n$ fluctuates in a manner which tends to become more markedly intermittent as n increases. This was confirmed very simply by taking photographs of oscillograph traces of u , $\partial u / \partial x$, $\partial^2 u / \partial x^2$ and $\partial^3 u / \partial x^3$, some of which are shown in

figure 6. (It should be noted that electrical noise is clearly visible in the records of $\partial^3 u / \partial x^3$, although since it has a uniform intensity the characteristics of the velocity variation are still discernible.) It is quite clear from these photographs that the higher derivatives fluctuate between limits which are sometimes large and sometimes small. Roughly speaking, the spatial field of turbulence is divided into a number of regions, in each of which there is a uniform activation of the higher wave-numbers but at a level of intensity which varies from region to region. The intensity within each region seems to tend to become either large or very small as the wave-number

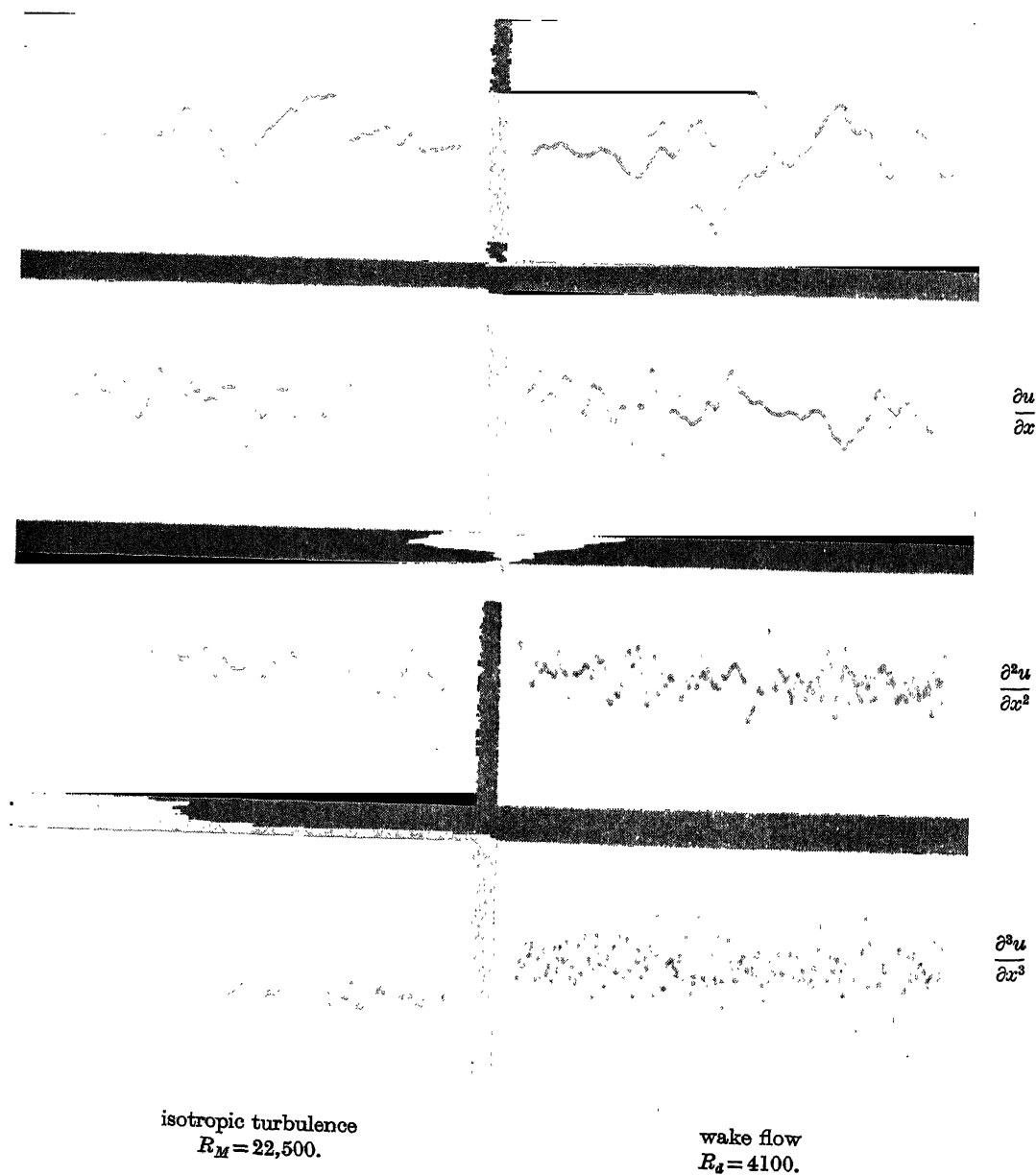


FIGURE 6. Oscillograms of fluctuations.

under discussion increases. The linear dimensions of the regions are large compared with the wave-lengths with which they are activated.

Finally, similar sets of measurements were made along the central plane of the turbulent wake of a circular cylinder in order to determine whether the statistical conditions associated with the equilibrium range of wave-numbers were affected in any way by the more rapid decay of the grid turbulence or by the presence of a mean velocity distribution. Two cylinder Reynolds numbers were used, 4100 and 680, and the technique was similar to that employed in the previous measurements. The measurements of $D_6 D_4^{-2} D_2$ are plotted in figure 7 in a form suitable for comparison with the earlier results for decaying isotropic turbulence, and all the points lie about a single curve as is predicted by the similarity theory. Likewise the measurements of the flattening factors of the velocity derivatives (included in figure 5) and the oscillograms of the velocity derivatives (included in figure 6) have features similar to the corresponding results for decaying isotropic turbulence. It can be inferred, first, that the form of the equilibrium spectrum (and its upper limit) is not dependent on statistical conditions outside the equilibrium range of wave-numbers, and secondly, that the spatial non-uniformity of the large wave-number components of the motion is likewise an intrinsic property of the equilibrium range of wave-numbers.

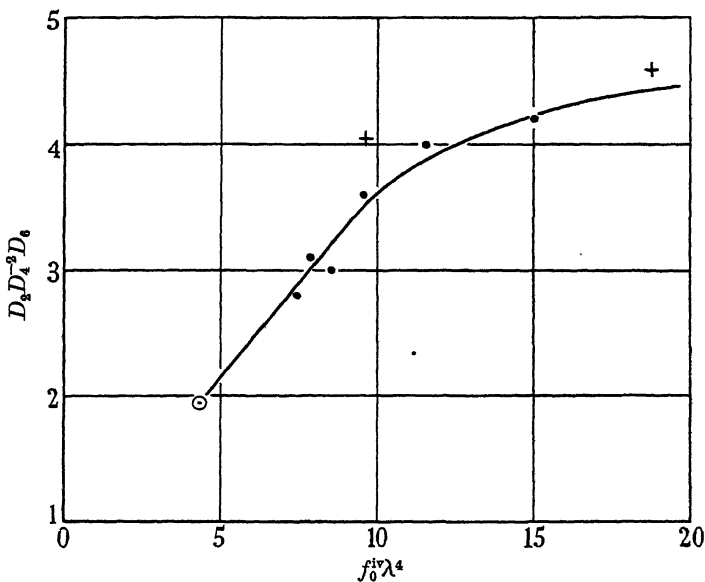


FIGURE 7. Variation of $D_6 D_4^{-2} D_2$ with $f_0^{iv} \lambda^4$.

•, isotropic turbulence; +, cylinder wake; ○, theoretical value at low Reynolds number.

5. DISCUSSION OF THE RESULTS

The measurements described in the preceding section indicate a state of affairs which has not been contemplated in previous theories of turbulent motion. The authors have no complete explanation to offer, and the following discussion is put forward as a guide to further thought on the subject. Certain negative conclusions at least are definite.

The basic observation which requires explanation is that activation of large wave-numbers is very unevenly distributed over space. These space variations in activation can be described as fluctuations in the spectrum at large wave-number. The measurements of flattening factors and the photographs of $\partial^n u / \partial x^n$ show clearly that the fluctuations are greater in magnitude for larger wave-numbers. As the wave-number is increased the fluctuations seem to tend to an approximate on-off, or intermittent, variation. Whatever the reason for the occurrence of these fluctuations, they appear to be intrinsic to the equilibrium range of wave-numbers. For it has been shown that $\alpha_n (n > 0)$ and $D_6 D_4^{-2} D_2$ are approximately invariant, first, with respect to Reynolds number of the grid, and secondly, with respect to the type of mean velocity distribution. All the evidence is consistent with the inference that the fluctuations are small in the region of the smallest wave-numbers of equilibrium range and become increasingly large at larger wave-numbers.

An immediate suggestion is that space is divided into a number of regions each of which has its own value of the energy dissipation ϵ and in each of which the energy spectrum has the universal form (1.8). Variations in $E(k)$ for a fixed value of k , from region to region, then depend on the corresponding variations in ϵ and it can be shown that under some circumstances the fluctuations in E are greater than those in ϵ . If E is proportional to ϵ^n , then

$$\frac{(\overline{E} - \bar{E})^2}{(\bar{E})^2} \approx n^2 \frac{(\overline{\epsilon} - \bar{\epsilon})^2}{(\bar{\epsilon})^2},$$

where bars represent averages over space. However, this suggestion is quite unable to account for the magnitude of the fluctuations in E and their rapid increase with k . For instance, when $E(k) \propto k^{-7}$, the corresponding value of n is 2, so that the (relative) standard deviation of E is only twice that of ϵ . In view of the evidence that the fluctuations arise within the equilibrium range, no finite value of n in the above relation is likely to give fluctuations in E as large as those observed.

It is natural to speculate that the high measured value of $D_6 D_4^{-2} D_2$ is in some way a consequence, or at least a concomitant, of the non-uniform spatial distribution of $E(k)$. Indeed, it was this speculation which led the authors to measure the flattening factors of $\partial^n u / \partial x^n$. However, no direct connexion can be traced. The quantity $D_6 D_4^{-2} D_2$ has an equal number of averages in the numerator and denominator so that it is not sensitive to intermittency. Provided that the measured quantities can be interpreted in the manner assumed by the *equilibrium theory*, the high measured value of $D_6 D_4^{-2} D_2$ must be regarded as evidence in favour of Heisenberg's expression for the spectrum and of the existence of an upper limit to the extent of the spectrum. Evidently the theoretical spectrum function $E(k)$, which has hitherto been assumed to be uniformly distributed in space (apart from small statistical fluctuations), corresponds in practice to the average value of the spectrum over a large region of space.

The authors suggest the following physical picture to account for the non-uniform distribution of the large wave-number components. First, consider how the non-uniformities may arise. It is well known in fluid dynamics that there is a strong tendency for a uniform distribution of vorticity to roll up into a number of strong

discrete vortices. For example, a vortex sheet or surface of velocity discontinuity is unstable to small disturbances and—which is more germane to the picture—soon degenerates into a distribution of isolated regions of concentrated vorticity. Evidently all other types of spatial distribution of vorticity are intrinsically unstable.* (Self-attracting matter has a similar property, although the interaction laws in the two cases are not identical.) It is natural to suppose that uniform distributions of the higher derivatives $\partial^2 u / \partial x^2$, $\partial^3 u / \partial x^3$, etc., are likewise unstable and tend to gather up into a number of discrete regions in which the intensity is high. Since this instability of uniform distributions of the velocity derivatives is intrinsic to the equilibrium range of wave-numbers, it will have a time-scale small compared with the time required for appreciable change in the total energy of the turbulence. Thus the instability will develop *rapidly* and will achieve a temporarily steady state, in which the large wave-number components are concentrated in isolated regions, early in the history of the turbulence.

Secondly, are the initial non-uniformities maintained throughout most of the decay—and if so, how?—or are they continually dying away and being re-formed? There is no direct evidence on this point, but the simpler hypothesis that the same non-uniformities persist throughout most of the decay seems plausible. The transfer of energy to a small range of wave-numbers depends in part on the whole system and in part on the energy already within that range. It is possible that a region of space in which $E(k)$ is above the equilibrium level will induce an above-average rate of transfer to energy to that region of space. This surplus of energy within a certain region of space and a certain range of wave-numbers will in turn lead to an above-average amount of energy associated with the next highest range of wave-numbers and lying within certain parts of that region of space.† Thus the process of subdivision of the fluid into regions of strong and of weak activation will occur in a stepwise manner as the wave-number—or the order of the derivative of the velocity—increases. The various regions in which wave-numbers of order k_n are activated will all lie within regions in which wave-numbers k_{n-1} of an order of magnitude smaller than k_n are activated, and these in turn are enclosed by regions in which even smaller wave-numbers are activated. Regions which are activated by wave-numbers of order k_n will probably compete with one another for the flow of energy from the smaller wave-numbers k_{n-1} , so that there will not be many such regions lying within each region activated by wave-numbers k_{n-1} . Consistent with this idea is the observation, from a study of the oscillograms of $\partial^3 u / \partial x^3$, that the mean separation of the visible activated regions is comparable with the integral scale of the turbulence, i.e. with the size of the energy-containing eddies.

Another statement of the whole process is that the distribution of vorticity (and of higher order derivatives) is *made* 'spotty' in the early stages of the decay by some intrinsic instability and is *kept* 'spotty' throughout the decay by the action of the

* The production of dissipation layers in the fluid, i.e. thin sheets inside which the velocity gradient is large, is suggested by the analysis of a mathematical model of turbulent motion made by Burgess (1948).

† It is possible that the existence of a lower critical Reynolds number, suggested by Heisenberg, assists in the maintenance of the peaks and hollows of the spatial distribution of the energy associated with large wave-numbers.

quadratic terms of the Navier-Stokes equations (e.g. by diffusive stretching of vortex lines). Apparently the rhyme originally due to L. F. Richardson,

Big whirls have little whirls,
That feed on their velocity;
And little whirls have lesser whirls,
And so on to viscosity.

must be taken more literally in turbulence than had been thought. The fact of transfer of energy from large-scale motions to small-scale motions has already been recognized in turbulence theory, and it now appears that the energy in the small-scale motion manifests itself as something approximating to a 'whirl', i.e. a region of high energy in the midst of a region of low energy.

APPENDIX. EFFECT OF DECAY OF TOTAL ENERGY ON THE UPPER LIMIT OF THE EQUILIBRIUM RANGE

In a lecture to the Seventh International Congress for Applied Mechanics, one of the authors (Batchelor 1948) used an argument, based on the idea of the time of reaction of a small range of wave-numbers, to show that the equilibrium spectrum—and in particular Heisenberg's k^{-7} law—could not extend to indefinitely large values of k in turbulence which is decaying. This argument can now be seen to be invalid, and, in view of the need for clear ideas about the possible reasons for an upper limit to the equilibrium range, will be corrected here. It was argued on dimensional grounds that the reaction time, or time required for energy transfer to occur, for any small range of wave-numbers dk is proportional to $[k^5 E(k)]^{-\frac{1}{2}} dk$, the implicit assumption being that k and $E(k)$ are the factors determining the local reaction time. The integrated total time for energy to traverse the whole equilibrium range then diverges when k is large (except in the very unlikely case that $E(k)$ decreases less rapidly than k^{-3} at large values of k), showing that, in decaying turbulence at least, there would not be sufficient time available for the whole of the equilibrium spectrum to be set up. The fallacy in the argument is that the time of reaction is *not* determined by the local spectrum quantities k and $E(k)$ any more than is the energy transfer itself.

However, a more accurate argument can be given. Let $S(k)$ be the rate of transfer of energy, owing to inertia effects, from wave-numbers lower to wave-numbers higher than k . Then the time $\tau(k)$ required for the transfer across wave-number k of sufficient energy to replenish the whole of the spectrum at wave-numbers greater than k is

$$\tau(k) = \int_k^\infty E(k) \frac{dk}{S(k)}.$$

If $\tau(k)$, for all values of k within the equilibrium range, is small compared with the time required for the over-all conditions (e.g. total energy) to change appreciably, then there is no objection on these grounds to an infinite upper limit to the equilibrium range. But for wave-numbers k within the equilibrium range we have

$$S(k) = 2\nu \int_k^\infty k^2 E(k) dk.$$

A necessary criterion for an infinite equilibrium range of wave-numbers in decaying turbulence is thus that

$$\tau(k) = \int_k^\infty E(k) dk / 2\nu \int_k^\infty k^2 E(k) dk \ll \frac{1}{\epsilon} \int_0^\infty E(k) dk$$

for indefinitely large values of k . For any wave-number within the equilibrium range, for which (1.8) is valid, we require

$$\int_\xi^\infty F(\xi) d\xi / \int_\xi^\infty \xi^2 F(\xi) d\xi \ll \frac{2}{\sqrt{(\nu\epsilon)}} \int_0^\infty E(k) dk,$$

where $\xi = k\nu^{\frac{1}{3}}\epsilon^{-\frac{1}{3}}$. The left side of this inequality has a single minimum value ξ_m^{-2} at $\xi = \xi_m$, where ξ_m is determined by

$$\xi_m^2 \int_{\xi_m}^\infty F(\xi) d\xi = \int_{\xi_m}^\infty \xi^2 F(\xi) d\xi.$$

For Heisenberg's spectrum (without cut-off), $\xi_m = \infty$, and the above condition is satisfied more strongly as k increases, so that there is sufficient time available, in decaying turbulence, for the setting up of an equilibrium spectrum of Heisenberg's form over an infinite range of wave-numbers. If the spectrum *does* terminate at some (large) wave-number, the reason for this cut-off appears not to be connected with the decay of total energy of the turbulence.

REFERENCES

- Batchelor, G. K. 1947 *Proc. Camb. Phil. Soc.* **43**, 533.
 Batchelor, G. K. 1948 *Proc. 7th Int. Congr. Appl. Mech., London, Introductory vol.*
 Batchelor, G. K. 1949 *Proc. Roy. Soc. A*, **195**, 513.
 Batchelor, G. K. & Townsend, A. A. 1947 *Proc. Roy. Soc. A*, **190**, 534.
 Batchelor, G. K. & Townsend, A. A. 1948 *Proc. Roy. Soc. A*, **193**, 539.
 Burgess, J. M. 1948 *Advances in App. Maths.*, **1**, 171.
 Heisenberg, W. 1948a *Z. Phys.* **124**, 628.
 Heisenberg, W. 1948b *Proc. Roy. Soc. A*, **195**, 402.
 Kolmogoroff, A. N. 1941 *C.R. Acad. Sci. U.R.S.S.* **30**, 301; **32**, 16.
 Kovasznay, L. S. G. 1948 *J. Aero. Sci.* **15**, 745.
 Obukhoff, A. 1941 *C.R. Acad. Sci. U.R.S.S.* **32**, 19.
 Taylor, G. I. 1938 *Proc. Roy. Soc. A*, **164**, 476.
 Townsend, A. A. 1947 *Proc. Camb. Phil. Soc.* **43**, 560.
 Townsend, A. A. 1948a *Proc. Camb. Phil. Soc.* **44**, 560.
 Townsend, A. A. 1948b *Aust. J. Sci. Res.* **1**, 161.

The analytic behaviour of Heisenberg's S matrix

By R. J. EDEN, *Peterhouse, University of Cambridge*

(Communicated by P. A. M. Dirac, F.R.S.—Received 25 May 1949)

The S matrix is considered for a system made up of an elementary particle being scattered on a fixed centre which has internal excited states. At threshold energy values for inelastic scattering, the S matrix undergoes abrupt changes of behaviour. A method of representing these as non-analytic changes in the matrix elements as functions of the total energy is suggested, and some of the implications investigated. It is shown that it may still be possible for the eigenvalues of S to be analytic functions of energy at the threshold values. The usual perturbation theory of quantum mechanics is used to consider a resonance scattering system of this type, and it is shown that the non-analytic changes in the matrix elements correspond to a non-analytic change in the unitary condition of S . When the incident particle is a photon, the excited states of the scatterer are necessarily unstable, and the S -matrix elements have singularities in the complex energy plane which correspond to these unstable levels. These singular points show clearly the connexion between the line widths for resonance scattering and Einstein's coefficients for spontaneous emission. It is shown that relative intensities of spectral lines may be obtained from the S matrix for scattering of light on an atom.

1. INTRODUCTION

In quantum mechanics collision cross-sections can be obtained directly from a certain unitary matrix S , which itself can be derived from a given Hamiltonian H . In deriving S from H , divergent integrals always occur and have to be eliminated by some subtraction procedure. Heisenberg (1943) has suggested that in a convergent theory the S matrix should replace the Hamiltonian as the fundamental quantity. Following a suggestion of Kramers, Heisenberg (1944) and Møller (1946) have shown, for a two-particle collision, that analytic continuation of the eigenvalues of S will lead to the energy levels of states in which the two particles are bound together. The method of analytic continuation has been extended by the author (1949) to systems of many particles, when the number of particles is a constant of the motion. When creation and annihilation of particles are allowed, the number of particles in a system is not an analytic function of the total energy. This causes the number of matrix elements in S to vary discontinuously with the energy, and one becomes doubtful of the applicability of the method of analytic continuation in the complex energy plane. Some study of the S matrix for a system with a varying number of particles has been made by Hu (1948*a*). Hu modifies the definition of the S matrix to make it an analytic function of energy. Unfortunately, this modification leads to a serious negative-energy difficulty, as Hu himself remarks; and it should be noted also that Hu's S matrix is not unitary. From a different viewpoint Dyson (1949) has studied the S matrix in quantum electrodynamics for a general system in which creation and annihilation of electron-positron pairs and of photons are permitted. Using the highly developed subtraction methods of Schwinger and Feynman, Dyson obtains S -matrix elements from a Hamiltonian; his results indicate that each matrix element of S will be non-analytic at threshold energies for particle creation, but one cannot conclude that this will be true for every type of interaction.

In this paper a simpler kind of discontinuous process is considered. This occurs when an incident particle is scattered on a fixed centre of force which itself has internal excited states of different energies. The centre can then absorb some of the kinetic energy of the particle and cause inelastic scattering. Such a system leads to several new aspects of S -matrix theory. Certain matrix elements of S , treated as functions of energy, have natural bounds at the threshold values for inelastic scattering. It is consistent with the physical interpretation to define these functions to be zero outside their natural bounds. These matrix elements are then non-analytic functions of energy, but it does not follow that every matrix element of S becomes non-analytic at each threshold value, or that the eigenvalues of S are non-analytic.

When the excited states of the scatterer, or force centre, are able to decay spontaneously, it may be of value to consider an approximate S matrix. If there exist states of absorption in which the particle is bound to the scatterer, resonance scattering may take place. Resonance reactions have been discussed by Hu (1948) from the viewpoint of one-particle scattering in S -matrix theory, and he shows that the theory provides a formal basis for such quantum-mechanical results as the Breit-Wigner formula for resonance dispersion. But the use of the one-particle viewpoint conceals the difficulties which arise when the scatterer has internal states of different energies. Here a form of S matrix is derived from quantum mechanics using Dirac's perturbation theory. The interaction is treated as a small perturbation, an approximation which is probably only valid when the incident particle is a photon. The matrix elements formally appear to be non-analytic, and it is shown how the unitary condition remains satisfied.

When the scattered particle is a photon certain new difficulties arise, since the states of absorption are then the same as the excited states of the scatterer. This leads to a connexion between the line widths for resonance dispersion of light on an atom, and the line widths for spontaneous emission. Both these phenomena are seen to result from singularities of the elastic scattering matrix elements of S in the complex energy plane. It is shown also that the S matrix can be used to determine the relative intensities of spectral lines.

Throughout this paper the word 'analytic' is used in a broad sense, usually implying that certain improper functions have been integrated out.

2. THE SCATTERING MATRICES ψ AND S

The notation and some of the results quoted are taken from Dirac (1948), but units are used in which $\hbar = c = 1$. The system considered is made up of a particle A of mass m and momentum $\mathbf{p}$, being scattered by another particle B . The scatterer B is assumed to be at rest, so that energy but not momentum has to be conserved; and B is taken to have internal excited states $|\alpha'\rangle$ of energy $H_S(\alpha')$. The kinetic energy of A is $W = \sqrt{(m^2 + p^2)}$, and the total energy of the system is a constant given by

$$E^0 = W^0 + H_S(\alpha^0), \quad (2.1)$$

where a superscript 0 denotes initial values.

A single scattering process consists of particle A coming from infinity, encountering particle B , and going off to infinity again. Since the total energy of the system is a constant, its state will be stationary; thus the classical picture of a particle moving through space is misleading, since the probability of any state is a constant in time. The reason for this is that the classical idea of the position of particle A has no meaning unless an observation is made to measure it. If this measurement is made, the system is disturbed and can no longer be considered in a stationary state. The essential feature of a stationary state, which will be used here, is that the probability of finding B in a particular state $|\alpha'\rangle$ is a constant. It is possible to form a classical picture of this quantum scattering process by considering a superposition in space of a large number of single scattering processes, each one being independent of the others. The incident wave can then be pictured as a steady stream of incoming particles, but it must be remembered that each particle is associated with its individual scatterer. Subject to limitation by the total energy E^0 , the scatterers B would then be divided between the various possible states $|\alpha^0\rangle, |\alpha'\rangle, \dots$, in a division which is independent of time.

If the system is prepared in an initial state $|\mathbf{p}^0\alpha^0\rangle$ with particle A incident, the wave function giving the probability that a second observation would show the system to be in a state $\langle\mathbf{p}\alpha'|\mathbf{p}^0\alpha^0\rangle$ can be written

$$\langle\mathbf{p}\alpha'|\psi|\mathbf{p}^0\alpha^0\rangle = \langle\mathbf{p}\alpha'|1|\mathbf{p}^0\alpha^0\rangle + \delta_+(W - W')\langle\mathbf{p}\alpha'|r|\mathbf{p}^0\alpha^0\rangle, \quad (2.2)$$

where

$$\delta_{\pm}(x) = \pm 1/(2\pi i x) + \frac{1}{2}\delta(x). \quad (2.3)$$

$W' = E^0 - H_S(\alpha')$ is the asymptotic value of the kinetic energy of the outgoing particle as required by energy conservation. It is the kinetic energy which the outgoing particle 'appears to have had' initially.

In writing the wave function in the form (2.2), the assumption is made that the states $|\alpha'\rangle$ and $|\alpha^0\rangle$ are stable. For simplicity, $\alpha^0, \alpha', \dots$ are taken to refer to ground level, first excited level, etc., of the scatterer B . Then $|\alpha^0\rangle$ is always stable. If $|\alpha'\rangle$ also is stable, or metastable, a wave function like (2.2) can be obtained in a stationary scattering process. However, the physical meaning of the state $|\mathbf{p}\alpha'\rangle$ is not very clear, since it refers to a measurement on the system when A is close to B , so that the kinetic energy W of A need not be conserved. Formally, a measurement of α is also made, and found to have the value α' . But experimentally neither W nor α is measured, only the asymptotic value W' is observed. If W' is less than the initial value W^0 by an amount $H_S(\alpha') - H_S(\alpha^0)$, it is explained by assuming the scatterer B has been left in an excited state $|\alpha'\rangle$. No such deduction can be made from an observed value of W when A is near to B . Thus the formalism referring to a state $|\mathbf{p}\alpha'\rangle$ has a clear meaning only in the asymptotic region where $W = W'$.

For all possible initial states $|\mathbf{p}^0\alpha^0\rangle$ having the same total energy E^0 , and all possible states $\langle\mathbf{p}\alpha'|$ in which energy need not be conserved, the wave functions (2.2) form the elements of the wave matrix ψ . These matrix elements are singular in the region of momentum space where energy is conserved. When a Hamiltonian H is known for the whole system the wave matrix ψ satisfies the operator equation

$$H\psi = E^0\psi. \quad (2.4)$$

If the interaction energy V can be regarded as an addition to the separate energies of A and B , $H = H_S + W + V$. Treating V as a small perturbation, it is possible to calculate ψ in various approximations; the type of approximation used depends on the aspect of the collision being considered. The wave function will still take the form (2.2) in an exact treatment.

The characteristic matrix S is defined by

$$\langle \mathbf{p}\alpha' | S | \mathbf{p}^0\alpha^0 \rangle = \langle \mathbf{p}\alpha' | 1 | \mathbf{p}^0\alpha^0 \rangle + \delta(W - W') \langle \mathbf{p}\alpha' | r | \mathbf{p}^0\alpha^0 \rangle. \quad (2.5)$$

It can be obtained at once from the form taken by the r matrix on the energy shell. A matrix will be described as 'on the energy shell' when the energies of the initial and final states of its matrix elements are the same; r will be used to denote such a matrix. The S matrix (2.5) is fully determined if one knows the asymptotic form of the ψ matrix for large separation of A and B . Thus the matrix S depends only on the region of co-ordinate space which is used for experimental observation. In particular, since in (2.5) the kinetic energy $W = \sqrt{(m^2 + p^2)} = W'$, there is a clear physical meaning to the state $\langle \mathbf{p}\alpha' |$.

It should be noted that the definition (2.5) does not correspond exactly to the original idea of defining the S matrix on the kinetic energy shell. The meaning has been extended by allowing the scatterer B to take up potential energy, that is, to jump into an excited state. But S is still determined entirely by the wave functions in the asymptotic region of co-ordinate space, and it is this that justifies the extension.

The quantity observed in practice is the differential cross-section for scattering into an angle $d\Omega$ about the direction $\mathbf{p}$. This is given by

$$\frac{W^0 W' p'}{p^0} |\langle \mathbf{p}\alpha' | r | \mathbf{p}^0\alpha^0 \rangle|^2 d\Omega, \quad (2.6)$$

where p denotes the length of the vector $\mathbf{p}$.

3. THRESHOLD ENERGIES FOR INELASTIC SCATTERING

In this section it is assumed that the scatterer B has only two states, a ground level denoted by $|\alpha^0\rangle$, and an excited level $|\alpha'\rangle$.

If the total energy of the system $E^0 = W^0 + H_S(\alpha^0)$, has a value less than $m + H_S(\alpha')$, energy conservation forbids the transition of the scatterer B to a final state $|\alpha'\rangle$; that is, inelastic scattering of particle A cannot occur. When E^0 is greater than $m + H_S(\alpha')$, inelastic scattering can occur, and the cross-section (2.6) for the transition $|\mathbf{p}\alpha^0\rangle$ to $\langle \mathbf{p}\alpha' |$ between states of equal energy is in general different from zero.

Let range (a) denote

$$m + H_S(\alpha^0) < E^0 < m + H_S(\alpha'), \quad (3.1a)$$

and range (b) denote

$$m + H_S(\alpha') < E^0. \quad (3.1b)$$

The value $E^0 = m + H_S(\alpha')$ is called the threshold energy for inelastic scattering. In a strict sense, the cross-section for inelastic scattering has no meaning in range

(a), but it is consistent with the physical interpretation to define it as zero in this range. This implies

$$\langle \mathbf{p}\alpha' | \underline{r} | \mathbf{p}^0\alpha^0 \rangle = 0, \quad \text{range (a),} \\ \neq 0, \quad \text{range (b),} \quad (3.2)$$

where $\underline{r}$ is the matrix on the energy shell.

Since the energy of the bra vector $\langle \mathbf{p}\alpha' |$ in real momentum space must be greater than $m + H_S(\alpha')$, the r matrix in range (a) must lie off the energy shell. The definition (3.2) gives a meaning to the $\underline{r}$ matrix elements outside the natural bound

$$E^0 = m + H_S(\alpha').$$

It is convenient to use the concept of states in 'complex momentum space'; representatives of these states will be distinguished by an asterisk. Thus the state $\langle * \mathbf{p}\alpha' |$ is allowed to have energy less than $m + H_S(\alpha')$. It is assumed that $\langle \mathbf{p}\alpha' | \underline{r} | \mathbf{p}^0\alpha^0 \rangle$ is an analytic function of the two variables E^0 and W in the ranges $E^0 > m + H_S(\alpha^0)$ and $W > m$. Then $\langle * \mathbf{p}\alpha' | \underline{r} | \mathbf{p}^0\alpha^0 \rangle$ is defined as the analytic continuation of $\langle \mathbf{p}\alpha' | \underline{r} | \mathbf{p}^0\alpha^0 \rangle$ as a function of E^0 and W . In particular, this will define $\langle * \mathbf{p}\alpha' | \underline{r} | \mathbf{p}^0\alpha^0 \rangle$ as a function different from zero in range (a). Note that in range (a)

$$| \mathbf{p}^0\alpha^0 * \rangle = | \mathbf{p}^0\alpha^0 \rangle.$$

Since in the wave function (2.2), $\langle \mathbf{p}\alpha' | \underline{r} | \mathbf{p}^0\alpha^0 \rangle$ is associated with a $\delta_+(W - W')$ factor, it will refer to an outgoing wave in the asymptotic region of co-ordinate space. This takes the form $\sim 1/r \exp(i pr)$ as $r \rightarrow \infty$. For large enough r , p is given by the conserved value $W = E^0 - H_S(\alpha')$; that is, the wave is determined by $\langle \mathbf{p}\alpha' | \underline{r} | \mathbf{p}^0\alpha^0 \rangle$. If analytic continuation is made to range (a) so that $W < m$, and $p = i |p|$, the form of the wave is $\sim 1/r \exp(-|p|r)$ in the asymptotic region of co-ordinate space, and represents a closed state. Thus $\langle * \mathbf{p}\alpha' | \underline{r} | \mathbf{p}^0\alpha^0 \rangle$ is associated in range (a) with a wave having exponential decay.

To obtain a representation of this wave $\sim 1/r \exp(-|p|r)$ in real momentum space, the usual Fourier transform can be taken. This will lead to a sum of Fourier components multiplying matrix elements of the type $\langle \mathbf{p}'\alpha' | \underline{r} | \mathbf{p}^0\alpha^0 \rangle$. In range (a) these matrix elements all lie off the energy shell. This shows that in range (a), $\langle * \mathbf{p}\alpha' | \underline{r} | \mathbf{p}^0\alpha^0 \rangle$ represents a sum of r matrix elements in real momentum space, all of which lie off the energy shell.

Corresponding to (3.2) the S matrix satisfies

$$\langle \mathbf{p}\alpha' | S | \mathbf{p}^0\alpha^0 \rangle = 0, \quad \text{in range (a),} \\ \neq 0 \quad \text{in range (b).} \quad (3.3)$$

The matrix elements of S become non-analytic functions of the total energy E^0 at the threshold value for inelastic scattering. This suggests the interesting question how the eigenvalues of S behave at threshold values. I do not think that this question can be fully answered at present, but it may be studied further by considering the behaviour of the elastic scattering matrix element $\langle \mathbf{p}\alpha^0 | S | \mathbf{p}^0\alpha^0 \rangle$ at the threshold value of E^0 .

The wave matrix element $\langle \mathbf{p}\alpha' | \psi | \mathbf{p}^0\alpha^0 \rangle$ need not lie on the energy shell so is not zero in range (a). However, its behaviour is essentially different for E^0 in ranges (a)

and (b); since for range (b) it has a singularity at $W = W' = E^0 - H_S(\alpha')$ due to the δ_+ factor, whilst for range (a) there is no such singularity. This function occurs as a factor in the integrand of the wave equation, which is

$$\int \langle \mathbf{p}\alpha^0 | V | \mathbf{p}''\alpha^0 \rangle d\mathbf{p}'' \langle \mathbf{p}''\alpha^0 | \psi | \mathbf{p}^0\alpha^0 \rangle + \int \langle \mathbf{p}\alpha^0 | V | \mathbf{p}''\alpha' \rangle d\mathbf{p}'' \langle \mathbf{p}''\alpha' | \psi | \mathbf{p}^0\alpha^0 \rangle \\ = (W^0 - W) \langle \mathbf{p}\alpha^0 | \psi | \mathbf{p}^0\alpha^0 \rangle. \quad (3.4)$$

This is the integral form of Schrödinger's equation. Normally one would seek as solution a wave function $\langle \mathbf{p}\alpha^0 | \psi | \mathbf{p}^0\alpha^0 \rangle$ which is analytic for all $W > m$ and $E^0 > m + H_S(\alpha^0)$. Whether it is possible to obtain such a solution depends on how far the form of the wave equation (3.4) changes as E^0 passes through the threshold value for inelastic scattering. An analytic function cannot satisfy two different equations for the ranges (a) and (b).

The critical term is the second integral in (3.4); its integrand contains a factor $\langle \mathbf{p}\alpha' | \psi | \mathbf{p}^0\alpha^0 \rangle$, which has a singularity on the path of integration for range (b) but not for range (a). It may happen that the interaction potential V is such that it compensates for this change, and the wave equation (3.4) has its form unchanged. Then it will be possible in general to find an analytic solution for $\langle \mathbf{p}\alpha^0 | \psi | \mathbf{p}^0\alpha^0 \rangle$. But if V does not compensate for the change in $\langle \mathbf{p}\alpha' | \psi | \mathbf{p}^0\alpha^0 \rangle$, there will be an essential change in the form of equation (3.4), and no analytic solution for $\langle \mathbf{p}\alpha^0 | \psi | \mathbf{p}^0\alpha^0 \rangle$ will exist. Two examples will be discussed which illustrate both possibilities.

$$\text{If} \quad \langle \mathbf{p}\alpha^0 | V | \mathbf{p}''\alpha' \rangle = \delta(W^a - W'') \langle \mathbf{p}\alpha^0 | v | \mathbf{p}''\alpha' \rangle, \quad (3.5)$$

$$\text{where} \quad W^a = W + H_S(\alpha^0) - H_S(\alpha'); \quad W'' = \sqrt{(m^2 + p''^2)}. \quad (3.6)$$

And $\delta(W^a - W'')$ is Dirac's δ function satisfying

$$\left. \begin{aligned} \int_a^\infty f(x) \delta(x - x_0) dx &= f(x_0) \quad \text{when } a < x_0, \\ &= 0 \quad \text{when } a > x_0. \end{aligned} \right\} \quad (3.7)$$

Only the second term in the wave equation (3.4) need be considered. When $W^a > m$, it is

$$\int \langle \mathbf{p}\alpha^0 | V | \mathbf{p}''\alpha' \rangle d\mathbf{p}'' \langle \mathbf{p}''\alpha' | \psi | \mathbf{p}^0\alpha^0 \rangle \\ = \iint \langle \mathbf{p}\alpha^0 | v | \mathbf{p}''\alpha' \rangle \delta(W^a - W'') \frac{dW'' dw''}{\sqrt{\Delta''}} \delta_+(W'' - W^b) \langle \mathbf{p}''\alpha' | r | \mathbf{p}^0\alpha^0 \rangle \quad (3.8)$$

$$= \delta_+(W - W^0) \int \langle \mathbf{p}\alpha^0 | v | \mathbf{p}''\alpha' \rangle \frac{dw''}{\sqrt{\Delta''}} \langle \mathbf{p}''\alpha' | r | \mathbf{p}^0\alpha^0 \rangle. \quad (3.9)$$

Where

$$W^b = W^0 + H_S(\alpha^0) - H_S(\alpha'),$$

$W''w''$ replace $\mathbf{p}''$, and

$$\Delta'' = \frac{\partial(W''w'')}{\partial(\mathbf{p}'')}. \quad (3.9)$$

In (3.9) W'' in the integrand is given the value W^a , i.e. the matrix v is on the energy shell.

When $W^a < m$, the wave equation (3.4) will not contain a term like (3.9). It follows that $\langle \mathbf{p}\alpha^0 | \psi | \mathbf{p}^0\alpha^0 \rangle$ will be a non-analytic function of W , at the point

$$W = m + H_S(\alpha') - H_S(\alpha^0).$$

Hence, since S is determined by the value for $W = W^0$, the elastic scattering matrix element $\langle \mathbf{p}\alpha^0 | S | \mathbf{p}^0\alpha^0 \rangle$ will be a non-analytic function of E^0 at the threshold value for elastic scattering.

One can obtain a class of interactions V , which give analytic behaviour, by a formal change in the above example. The Dirac δ function in V is replaced by the Heisenberg-Møller δ function, which satisfies

$$\int_a^\infty f(x) \delta(x, x_0) dx = f(x_0) \quad \text{for all } a. \quad (3.10)$$

The properties of this δ function are given by Møller (1946); it is a function of two variables, and the notation $\delta(x, x_0)$ distinguishes it from the Dirac δ function $\delta(x - x_0)$, which is a function of the one real variable $(x - x_0)$ only. Implicit in this change is the use of complex momentum space in the integrands of (3.4); the interpretation of this change is considered below. It gives

$$\langle \mathbf{p}\alpha^0 | V | \mathbf{p}''\alpha' \rangle = \delta(W^a, W'') \langle \mathbf{p}\alpha^0 | v | \mathbf{p}''\alpha' \rangle, \quad (3.11)$$

and the second integral of (3.4) becomes

$$\delta_+(W - W^0) \int \langle \mathbf{p}\alpha^0 | v | \mathbf{p}''\alpha' * \rangle \frac{dw''}{\sqrt{\Delta''}} \langle * \mathbf{p}''\alpha' | r | \mathbf{p}^0\alpha^0 \rangle, \quad (3.12)$$

where $W'' = W^a$ in the integrand. The term (3.12) is present in the wave equation for all values of $W > m$. Thus the wave equation is analytic, and one would expect to find an analytic wave function $\langle \mathbf{p}\alpha^0 | \psi | \mathbf{p}^0\alpha^0 \rangle$ as a solution.

The formal change which leads to this analytic behaviour is not so artificial as it might at first appear. For $W'' \doteq W^a < m$, the matrix element $\langle * \mathbf{p}''\alpha' | r | \mathbf{p}^0\alpha^0 \rangle$ corresponds to a wave $\sim 1/r \exp(-|p''|r)$ for large r in co-ordinate space. So the interaction $\langle \mathbf{p}\alpha^0 | v | \mathbf{p}''\alpha' * \rangle$ means that a transition can take place to a virtual state in which the particle A appears to be bound, while the scatterer B is excited to $|\alpha'\rangle$. The total energy in this picture appears to be conserved and equal to E^0 . Alternatively, if measurements of the momentum of particle A were made in the intermediate state, they would correspond to interaction matrix elements $\langle \mathbf{p}\alpha^0 | v | \mathbf{p}''\alpha' \rangle$ off the energy shell in real momentum space. In this picture the interaction V is such that when W (in (3.4)) drops below $m + H_S(\alpha') - H_S(\alpha^0)$, matrix elements of V off the energy shell come into play, and compensate exactly the change in behaviour of $\langle \mathbf{p}\alpha' | \psi | \mathbf{p}^0\alpha^0 \rangle$.

This gives an analytic function $\langle \mathbf{p}\alpha^0 | \psi | \mathbf{p}^0\alpha^0 \rangle$ throughout the ranges (a) and (b). Hence the elastic scattering matrix element $\langle \mathbf{p}\alpha^0 | S | \mathbf{p}^0\alpha^0 \rangle$ is an analytic function of E^0 . A special example of this type has been solved by Touschek (1947). He works in co-ordinate space of one dimension only, and makes the basic assumption that the wave equation (the non-relativistic Schrödinger equation) does not change its form at the threshold energy value. In co-ordinate representation the interaction

energy takes quite a simple form. An explicit solution is obtained, which shows that the wave function contains a term having a branch point in the complex plane at the threshold value.

The results of this section show that the inelastic scattering matrix elements of S can be considered as functions of the total energy which become non-analytic at threshold values. The elastic scattering matrix element may, or may not, be analytic, depending on the type of interaction potential. In the next section, the investigation will be continued admitting both possibilities.

4. THE EIGENVALUES OF S AT THRESHOLD ENERGIES

The eigenvalue equation is

$$S\phi = \lambda\phi. \quad (4.1)$$

Since the total energy commutes with S , λ is a function of E^0 . For range (a), the S matrix can be written

$$S = S_a = \begin{pmatrix} \langle \mathbf{p}\alpha^0 | S_a | \mathbf{p}^0\alpha^0 \rangle & 0 \\ 0 & - \end{pmatrix}. \quad (4.2)$$

$\langle \mathbf{p}\alpha' | S_a | \mathbf{p}'\alpha' \rangle$ will not be required explicitly, so need not be defined in this range. For range (b),

$$S = S_b = \begin{pmatrix} \langle \mathbf{p}\alpha^0 | S_b | \mathbf{p}^0\alpha^0 \rangle & \langle \mathbf{p}\alpha^0 | S_b | \mathbf{p}'\alpha' \rangle \\ \langle \mathbf{p}\alpha' | S_b | \mathbf{p}^0\alpha^0 \rangle & \langle \mathbf{p}\alpha' | S_b | \mathbf{p}'\alpha' \rangle \end{pmatrix}. \quad (4.3)$$

In (4.3), $W^0 + H_S(\alpha^0) = E^0 = W' + H_S(\alpha')$, and S contains a Dirac δ factor in total energy. Both (4.2) and (4.3) satisfy the unitary condition

$$SS^\dagger = S^\dagger S = 1. \quad (4.4)$$

Hence, if λ_a^0 is an eigenvalue of (4.2) satisfying

$$\int \langle \mathbf{p}\alpha^0 | S_a | \mathbf{p}''\alpha^0 \rangle d\mathbf{p}'' \langle \mathbf{p}''\alpha^0 | \phi \rangle = \lambda_a^0 \langle \mathbf{p}\alpha^0 | \phi \rangle, \quad (4.5)$$

and λ_b is an eigenvalue of (4.3) satisfying

$$\int \langle \mathbf{p}\alpha^0 | S_b | \mathbf{p}''\alpha^0 \rangle d\mathbf{p}'' \langle \mathbf{p}''\alpha^0 | \phi \rangle + \int \langle \mathbf{p}\alpha^0 | S_b | \mathbf{p}''\alpha' \rangle d\mathbf{p}'' \langle \mathbf{p}''\alpha' | \phi \rangle = \lambda_b \langle \mathbf{p}\alpha^0 | \phi \rangle, \quad (4.6)$$

then

$$\begin{cases} \lambda_a^0 \lambda_a^{0*} = 1 & \text{in range (a),} \\ \lambda_b \lambda_b^* = 1 & \text{in range (b).} \end{cases} \quad (4.7)$$

The question to be answered is whether λ_a^0 and λ_b can be the same analytic function of E^0 .

If the wave function $\langle \mathbf{p}\alpha^0 | \psi | \mathbf{p}^0\alpha^0 \rangle$ is an analytic function of E^0 at $m + H_S(\alpha')$, it follows that $\langle \mathbf{p}\alpha^0 | S | \mathbf{p}^0\alpha^0 \rangle$ is also analytic. The eigenvalues λ_a^0 are determined by this term in energy range (a). If $\langle \mathbf{p}\alpha^0 | S | \mathbf{p}^0\alpha^0 \rangle$ is analytic, it will also determine the same eigenvalue function λ_a^0 in range (b). Now S_a is unitary in range (a), and S_b is unitary in range (b). Moreover, $\langle \mathbf{p}\alpha^0 | S_a | \mathbf{p}^0\alpha^0 \rangle$ and $\langle \mathbf{p}\alpha^0 | S_b | \mathbf{p}^0\alpha^0 \rangle$ are assumed to be the same analytic function of E^0 . Hence, in general, it is not possible for S_a and S_b to be unitary over the same energy range. In particular, S_a will not be

unitary in range (b), so that $|\lambda_a^0| \neq 1$ in range (b). But $|\lambda_b| = 1$ in range (b), hence λ_a^0 and λ_b cannot be the same analytic function of E^0 . Thus analytic behaviour of the elastic scattering matrix element of S , in general, implies that the eigenvalues of S are not analytic at the threshold energies, and conversely.

If $\langle \mathbf{p}\alpha^0 | S | \mathbf{p}^0\alpha^0 \rangle$ is not analytic, it is possible for S_a to be unitary not only in range (a) but also in range (b). If this is so, then $|\lambda_a^0| = 1$ for both (a) and (b). But it is still not consistent to take λ_a^0 and λ_b to be the same function of E^0 . This is because the matrix S_b has more eigenvalues than are given by λ_a^0 , which arise solely from the first term of S_a . It is consistent, however, to take half the eigenvalues, λ_b^0 say, of λ_b to be the same functions as λ_a^0 . Denote the remaining eigenvalues by λ'_b . Then $|\lambda_b^0| = |\lambda_a^0| = 1$ in both (a) and (b), but it is not essential that $|\lambda'_b| = 1$ in range (a).

This shows that the following assumption may be regarded as an alternative to the assumption of analytic behaviour of the elastic scattering matrix element: The functions of energy λ_a^0 , which are the eigenvalues of S in range (a), are included amongst the eigenvalues of S in range (b). This is not the only alternative assumption, but it appears to be the simplest.

There is a trivial case where it is possible for both assumptions to hold. This occurs when the system only gives 'pure scattering', i.e. when the incident particle A cannot cause transitions in the scatterer B . Then for range (b),

$$S = \begin{pmatrix} \langle \mathbf{p}\alpha^0 | S | \mathbf{p}^0\alpha^0 \rangle & 0 \\ 0 & \langle \mathbf{p}\alpha' | S | \mathbf{p}'\alpha' \rangle \end{pmatrix}. \quad (4.8)$$

The functions λ_b^0 are eigenvalues of $\langle \mathbf{p}\alpha^0 | S | \mathbf{p}^0\alpha^0 \rangle$ and therefore the same as λ_a^0 , while λ'_b are eigenvalues of $\langle \mathbf{p}\alpha' | S | \mathbf{p}'\alpha' \rangle$. One has to suppose that there is some other method of finding the state of B than by observing with particle A .

5. SCATTERER WITH UNSTABLE EXCITED STATES

In some important practical cases, the excited states $|\alpha'\rangle$ of the scatterer B are unstable. For example, if B is an atom, it can usually decay from one excited state to another, or to its ground state, by spontaneous emission of a photon. It is assumed now that the excited states $|\alpha'\rangle$, $|\alpha''\rangle$, ... of B can decay spontaneously by emission of a particle C of momentum $\mathbf{q}$. This appears to raise a new difficulty in S -matrix theory, since the levels of $|\alpha'\rangle$, etc., will now have a band width in the energy spectrum. It is assumed at first that this width is sufficiently narrow to justify speaking of a particular value $H_S(\alpha')$ as the energy of the level $|\alpha'\rangle$, but it is shown later that such an assumption is not essential to fix a specific value for the threshold energy.

It is now necessary to consider carefully the implications of a steady scattering process in quantum mechanics. In the classical picture, a single collision consists of particle A coming from infinity, encountering particle B (initially in the ground state $|\alpha^0\rangle$), and going off to infinity again. B may be left in the ground state, in which case the collision is complete. Or, if $E^0 > m + H_S(\alpha')$, B may be left in an excited state $|\alpha'\rangle$, in which case there is inelastic scattering of particle A . However,

the collision is not complete. At some subsequent time B will decay from $|\alpha'\rangle$ to $|\alpha^0\rangle$ with emission of particle C , which then goes off to infinity and the collision is complete.

In the quantum picture of a steady scattering process, the probability of any state of the system is constant in time. The concept of a lapse of time before emission of particle C does not arise. The observed results, from an initial state $|\mathbf{p}^0\alpha^0\rangle$, will be a certain probability for a final state $\langle\mathbf{p}\alpha^0|$, and a certain probability for a final state $\langle\mathbf{p}\mathbf{q}\alpha^0|$. Both final states have energy E^0 . The observed result in the second case is the two outgoing particles $\mathbf{p}$ and $\mathbf{q}$. In a strictly stationary process no more information is obtained, and it has only a formal meaning to say that excited states $|\alpha'\rangle$ exist. The observed result is solely the creation of a new particle C with momentum $\mathbf{q}$. When C has a finite rest mass, the threshold energy for particle creation is given by $E^0 = H_S(\alpha^0) + m_A + m_C$, where m_A and m_C are the rest masses of particles A and C . Providing m_A and m_C have no band spread, the threshold energy value is precise.

The matrix elements of ψ are of the following two types:

$$\langle\mathbf{p}\alpha^0|\psi|\mathbf{p}^0\alpha^0\rangle \quad \text{and} \quad \langle\mathbf{p}\mathbf{q}\alpha^0|\psi|\mathbf{p}^0\alpha^0\rangle. \quad (5.1)$$

Corresponding to (3.3),

$$\begin{aligned} \langle\mathbf{p}\mathbf{q}\alpha^0|S|\mathbf{p}^0\alpha^0\rangle &= 0 & \text{for } E^0 < H_S(\alpha^0) + m_A + m_C, \\ &\neq 0 & \text{for } E^0 > H_S(\alpha^0) + m_A + m_C. \end{aligned} \quad (5.2)$$

The matrix elements of S are non-analytic at the threshold energy for particle creation, and the considerations of §§ 3 and 4 apply here also. It should be noted that the threshold energy for particle creation corresponds to the lower end of the band width of the unstable level $H_S(\alpha')$. Hence

$$H_S(\alpha^0) + m_C < H_S(\alpha'). \quad (5.3)$$

The uncertainty of the energy $H_S(\alpha')$ does not lead to any uncertainty in the point where the S matrix elements become non-analytic.

If the decay lifetime of an excited state $|\alpha'\rangle$ of B is long compared with the time spent by particle A near to B , it is possible to consider an experiment which leads to inelastic scattering of A without creation of C . Such a result would be obtained when the incident wave has a sharp beginning. Then during the initial stages of the collision, particle A may be observed inelastically scattered, but the probability of finding a particle C created would be very small. To a good approximation, the experiment would be observing a transition $|\mathbf{p}\alpha^0\rangle$ to $\langle\mathbf{p}\alpha'|$ of equal energy, $E^0 > H_S(\alpha') + m_A$. In this case it may be of value to consider an approximate S matrix, which can be written in the form (4.2) for energy range (a), and in the form (4.3) for energy range (b). The condition for the validity of such an approximation is the same as the condition that the transitions described by it are observable, namely that the decay period of $|\alpha'\rangle$ is long compared with the duration of the collision. It should be noted that this approximate S matrix is non-analytic at $E^0 = H_S(\alpha') + m_A$, and this may be very different from the point $E^0 = H_S(\alpha^0) + m_A + m_C$, at which the exact S matrix is non-analytic.

If, as suggested by Heisenberg (1949), the rest masses of elementary particles have a band width, then the S matrix itself must be regarded as an approximate concept, but such a possibility would lead to speculations beyond the scope of this paper.

6. RESONANCE SCATTERING

In this section the perturbation method of quantum mechanics is used to form an S matrix. This method is perhaps only applicable when the incident particle is a photon, but it gives interesting confirmation of the S -matrix theory of unstable states, and throws more light on the non-analytic behaviour. The system considered is formed from an incident particle A , mass m , kinetic energy W , and momentum $\mathbf{p}$, and a scatterer B with stable excited states $|\alpha'\rangle$ of energy $H_S(\alpha')$. There exist states $|k\rangle$ of energy E_k for which particle A is absorbed in the scatterer B . The total energy of the system is $E^0 = W^0 + H_S(\alpha^0)$, and the interaction V between A and B is treated as a small perturbation.

The wave matrix is given by (2.2), and the S matrix by (2.5),

$$\langle \mathbf{p}\alpha' | S | \mathbf{p}^0\alpha^0 \rangle = \langle \mathbf{p}\alpha' | 1 | \mathbf{p}^0\alpha^0 \rangle + \delta(W - W') \langle \mathbf{p}\alpha' | r | \mathbf{p}^0\alpha^0 \rangle,$$

where $W' = E^0 - H_S(\alpha')$ and $E^0 > H_S(\alpha') + m$. When $E^0 \doteq E_k$,

$$\langle \mathbf{p}\alpha' | r | \mathbf{p}^0\alpha^0 \rangle \doteq -2\pi i \frac{\langle \mathbf{p}\alpha' | V | k \rangle \langle k | V | \mathbf{p}^0\alpha^0 \rangle}{(E^0 - E_k - a + ib)}. \quad (6.1)$$

Only the value r of r on the energy shell is important. In (6.1) a and b are real, and b is positive. They are

$$a = \sum_{\alpha'} \int |\langle k | V | \mathbf{p}\alpha' \rangle|^2 d\mathbf{p} / (W' - W), \quad (6.2)$$

$$b = \pi \sum_{\alpha'} \int |\langle k | V | \mathbf{p}\alpha' \rangle|^2 d\mathbf{p} \delta(W' - W), \quad (6.3)$$

where $\sum_{\alpha'}$ is taken only over those values of α' for which $W' > m$.

Variation of E^0 in the complex plane will lead to the same singularities in r as in S . Hence the matrix elements of S have a singularity at

$$E^0 = E_k + a - ib. \quad (6.4)$$

According to the general theory of the S matrix, this singularity, which is in the lower half complex E^0 plane, means that the state $|k\rangle$ is unstable and has decay constant $2b$. This interpretation is confirmed by a direct calculation of the decay probability of $|k\rangle$ (Dirac 1948).

The mean energy of $|k\rangle$ appears to be $E_k + a$, and not E_k as was first assumed. The interpretation of this is quite straightforward. The value $E_k + a$, as given by S -matrix theory, is the observable value of the mean energy and is obtained either by the resonance point for scattering, or by observing the maximum intensity of particles when $|k\rangle$ decays. But E_k is not an observable, it is an auxiliary quantity introduced for convenience in calculation. Formally, E_k is the energy of the undisturbed state $|k\rangle$, in which particle A is bound to the scatterer B . But as soon as an observation is made to obtain the energy value, the state is disturbed and falls into an eigenstate of the observation. $E_k + a$ is the eigenvalue of this observation,

but E_k is the 'energy' of state $|k\rangle$ and is not a well-defined concept in quantum mechanics.

In addition to matrix elements like (6.1), the r matrix will contain elements for absorption. When $E^0 \doteq E_k$, these are

$$\langle k | r | \mathbf{p}^0 \alpha^0 \rangle = -2\pi i \frac{\langle k | V | \mathbf{p}^0 \alpha^0 \rangle}{E^0 - E_k - a + ib}. \quad (6.5)$$

These lead to the wave function $\langle k | \psi | \mathbf{p}^0 \alpha^0 \rangle$, giving the probability that a 'measurement' would reveal a state of absorption. In a steady scattering process the probability of the state $|k\rangle$ must be constant in time. Therefore the formation probability must be exactly balanced by the decay probability. In a Gibbs ensemble picture of a large number of individual collisions superimposed, one can say the number of particles A absorbed is exactly equal to the number emitted. Only when a non-stationary process is considered can an absorption be observed, so in an exact treatment the S matrix contains no elements corresponding to (6.5).

When E^0 is not near to a resonance value E_k , the r matrix is given by the dispersion approximation in quantum mechanics. It is

$$\langle \mathbf{p} \alpha' | r | \mathbf{p}^0 \alpha^0 \rangle = -2\pi i \left\{ \langle \mathbf{p} \alpha' | V | \mathbf{p}^0 \alpha^0 \rangle + \sum_{k'} \frac{\langle \mathbf{p} \alpha' | V | k' \rangle \langle k' | V | \mathbf{p}^0 \alpha^0 \rangle}{E^0 - E_{k'}} \right\}. \quad (6.6)$$

This assumes that $\langle \mathbf{p} \alpha' | V | \mathbf{p}^0 \alpha^0 \rangle$ is second order in smallness, when $\langle \mathbf{p} \alpha' | V | k \rangle$ is first order.

Formulae (6.1) and (6.6) will now be used to investigate how the unitary condition, $S^\dagger S = S S^\dagger = 1$, can hold at a threshold value where the S matrix becomes non-analytic. Writing $S = 1 + R$, the unitary condition is

$$R + R^\dagger + R R^\dagger = 1, \quad R R^\dagger = R^\dagger R. \quad (6.7)$$

If $m + H_S(\alpha')$ is not near to E_k , the dispersion approximation (6.6) has to be used. This gives

$$\langle \mathbf{p}^0 \alpha^0 | r | \mathbf{p} \alpha' \rangle_2^* = 2\pi i \left\{ \langle \mathbf{p} \alpha' | V | \mathbf{p}^0 \alpha^0 \rangle + \sum_k \frac{\langle \mathbf{p} \alpha' | V | k \rangle \langle k | V | \mathbf{p}^0 \alpha^0 \rangle}{E^0 - E_k} \right\}, \quad (6.8)$$

where a subscript 2 denotes a second-order approximation. Then

$$\langle \mathbf{p} \alpha' | R | \mathbf{p}^0 \alpha^0 \rangle = \delta(W - W') \langle \mathbf{p} \alpha' | r | \mathbf{p}^0 \alpha^0 \rangle, \quad (6.9)$$

$$\langle \mathbf{p} \alpha' | R^\dagger | \mathbf{p}^0 \alpha^0 \rangle = \delta(W - W') \langle \mathbf{p}^0 \alpha^0 | r | \mathbf{p} \alpha' \rangle^*. \quad (6.10)$$

The fourth-order approximation for r gives, for r on the energy shell,

$$\begin{aligned} \langle \mathbf{p} \alpha' | r | \mathbf{p}^0 \alpha^0 \rangle_4 = & -4\pi^2 \left\{ \sum_{\alpha''} \int \langle \mathbf{p} \alpha' | V | \mathbf{p}'' \alpha'' \rangle d\mathbf{p}'' \langle \mathbf{p}'' \alpha'' | V | \mathbf{p}^0 \alpha^0 \rangle \delta_+(E'' - E^0) \right. \\ & + \sum_{\alpha''} \int \sum_k \frac{\langle \mathbf{p} \alpha' | V | \mathbf{p}'' \alpha'' \rangle d\mathbf{p}'' \langle \mathbf{p}'' \alpha'' | V | k \rangle \langle k | V | \mathbf{p}^0 \alpha^0 \rangle \delta_+(E'' - E^0)}{(E'' - E_k)} \\ & + \sum_{\alpha''} \int \sum_k \frac{\langle \mathbf{p} \alpha'' | V | k \rangle \langle k | V | \mathbf{p}'' \alpha'' \rangle d\mathbf{p}'' \langle \mathbf{p}'' \alpha'' | V | \mathbf{p}^0 \alpha^0 \rangle \delta_+(E'' - E^0)}{(E^0 - E_k)} \\ & \left. + \sum_{\alpha''} \int \sum_k \sum_{k'} \frac{\langle \mathbf{p} \alpha'' | V | k \rangle \langle k | V | \mathbf{p}'' \alpha'' \rangle d\mathbf{p}'' \langle \mathbf{p}'' \alpha'' | V | k' \rangle \langle k' | V | \mathbf{p}^0 \alpha^0 \rangle \delta_+(E'' - E^0)}{(E^0 - E_k)(E'' - E_{k'})} \right\}, \end{aligned} \quad (6.11)$$

where $E'' = W'' + H_S(\alpha'')$, and $\sum_{\alpha''}$ is only over states with $E^0 > m + H_S(\alpha'')$. By (6.6) and (6.8) to (6.11), $R_2^\dagger = -R_2$, hence

$$R_2^\dagger R_2 = R_2 R_2^\dagger, \quad (6.12)$$

and using (2.3),

$$R_4 + R_4^\dagger + R_2 R_2^\dagger = 0. \quad (6.13)$$

This shows that the unitary condition holds to the fourth approximation. The abrupt change in the inelastic scattering elements of the S matrix, as E^0 increases through $m + H_S(\alpha'')$, leads to an extra term in the unitary equation (6.13), which is

$$\int \langle \mathbf{p}\alpha^0 | R_2 | \mathbf{p}''\alpha'' \rangle d\mathbf{p}'' \langle \mathbf{p}''\alpha'' | R_2^\dagger | \mathbf{p}^0\alpha^0 \rangle. \quad (6.14)$$

The term (6.14) is, in this instance, cancelled by an extra term appearing in R_4 and $R_4^\dagger$. Thus the unitary condition depends on a non-analytic change in the elastic scattering matrix elements of S , at the threshold energies for inelastic scattering.

If $E^0 \doteq E_k \doteq m + H_S(\alpha'')$, the resonance scattering approximation (6.1) must be used. This gives

$$\langle \mathbf{p}\alpha' | r^\dagger | \mathbf{p}^0\alpha^0 \rangle = 2\pi i \frac{\langle \mathbf{p}\alpha' | V | k \rangle \langle k | V | \mathbf{p}^0\alpha^0 \rangle}{(E^0 - E_k - a - ib)}. \quad (6.15)$$

It should be noted that b , given by (6.3), like R_4 (6.11), is not an analytic function of energy at $E^0 = m + H_S(\alpha'')$, due to $\sum_{\alpha''}$ being taken only over states with

$$E^0 > m + H_S(\alpha'').$$

From (6.1) and (6.15),

$$\langle \mathbf{p}\alpha' | r + r^\dagger | \mathbf{p}^0\alpha^0 \rangle = -4\pi b \frac{\langle \mathbf{p}\alpha' | V | k \rangle \langle k | V | \mathbf{p}^0\alpha^0 \rangle}{(E^0 - E_k - a - ib)(E^0 - E_k - a + ib)}. \quad (6.16)$$

Substitute in (6.16) the value given by (6.3), and use (6.9) and (6.10). Then

$$\begin{aligned} \langle \mathbf{p}\alpha^0 | R + R^\dagger | \mathbf{p}^0\alpha^0 \rangle &= -4\pi^2 \delta(W - W^0) \sum_{\alpha''} \int \frac{\langle \mathbf{p}\alpha^0 | V | k \rangle \langle k | V | \mathbf{p}''\alpha'' \rangle \delta(E'' - E^0) d\mathbf{p}''}{(E^0 - E_k - a - ib)} \\ &\quad \times \frac{\langle \mathbf{p}''\alpha'' | V | k \rangle \langle k | V | \mathbf{p}^0\alpha^0 \rangle}{(E^0 - E_k - a + ib)} \end{aligned} \quad (6.17)$$

$$= -\langle \mathbf{p}\alpha^0 | R^\dagger R | \mathbf{p}^0\alpha^0 \rangle. \quad (6.18)$$

It is clear also that $RR^\dagger = R^\dagger R$, so the unitary condition holds. The extra term in the product (6.18), as E^0 increases above $m + H_S(\alpha'')$, cancels with an extra term in the series for b . Thus it is again the non-analytic behaviour of the elastic scattering matrix elements which ensures the unitary condition.

These calculations from usual quantum mechanics indicate that at threshold values for new real processes, where the S matrix undergoes an abrupt change of behaviour, every matrix element of S will take a different analytic form. But it should be noted that the interaction considered here corresponds to that in equations (3.5) to (3.9). It is not mathematically necessary that a change in form of the unitary equation corresponds to a change in analytic behaviour of the elastic scattering elements, since the operation of taking the conjugate complex of a function is not analytic, in particular it is affected by branch points in the complex energy plane.

7. PHOTON SCATTERING

When the scattered particle A is a photon, the states of absorption $|k\rangle$ are the same as the excited states $|\alpha''\rangle$ of the scatterer B . This means that these excited states are unstable, and the decay particle C is a photon. Since the rest mass of C is now zero, inelastic scattering—or more exactly creation of a photon pair—is possible for all energies $E^0 > H_S(\alpha^0)$. This has the consequence that the elastic scattering matrix elements must be analytic at the energies $H_S(\alpha'')$. It will be verified that this result is given by Dirac's perturbation theory, and a connexion with coefficients for spontaneous emission will be noted.

Denote the momentum of the photon by $\mathbf{p}$, its polarization by $\mathbf{l}$, and frequency by ν . Then since $\hbar = c = 1$, $|p| = 2\pi\nu$. The matrix elements for interaction with a one-electron atom are given by Dirac (1948, § 63), and in resonance-scattering approximation lead to

$$\left. \begin{aligned} \langle \mathbf{p} \mathbf{l} \alpha^0 | r | \mathbf{p}^0 \mathbf{l}^0 \alpha^0 \rangle &= -e^2 \frac{\nu(\alpha'' \alpha^0) \nu(\alpha'' \alpha^0)}{(\nu \nu^0)^{\frac{1}{2}}} \frac{\langle \alpha^0 | x_1 | \alpha'' \rangle \langle \alpha'' | x_{1^0} | \alpha^0 \rangle}{E^0 - H_S(\alpha'') - a + ib}, \\ 2\pi\nu(\alpha'' \alpha^0) &= H_S(\alpha'') - H_S(\alpha^0), \quad 2\pi\nu^0 = |p^0|. \end{aligned} \right\} \quad (7.1)$$

x_{1^0} is the component of electron co-ordinate in the direction $\mathbf{l}^0$ of the photon $\mathbf{p}^0$. Energy conservation on going over to the r or S matrix requires $\nu = \nu^0$. The line shift a , and the line width $2b$, are given by

$$a = 2\pi e^2 \sum_{\alpha'} (\nu(\alpha'' \alpha'))^2 \iint \sum_1 |\langle \alpha'' | x_1 | \alpha' \rangle|^2 \frac{\nu d\nu \sin \omega d\omega d\chi}{(\nu' - \nu)}, \quad (7.2)$$

$$b = 2\pi^2 e^2 \sum_{\alpha'} (\nu(\alpha'' \alpha'))^2 \iint \nu' \sum_1 |\langle \alpha'' | x_1 | \alpha' \rangle|^2 \sin \omega d\omega d\chi, \quad (7.3)$$

where $2\pi\nu' = E^0 - H_S(\alpha')$, and $\sum_{\alpha'}$ is taken only over α' such that $\nu' > 0$.

The term (7.2) giving the shift of the energy level diverges. Until recently this term has been ignored, but it has been shown by Bethe that a subtraction method consistent with relativity requirements will remove all but a finite part a_1 of a . This finite part a_1 is interpreted as the shift in the energy level first observed in hydrogen by Lamb and Retherford. The imaginary term ib in the singularity of (7.1), shows that the excited state $|\alpha''\rangle$ of the atom can decay by spontaneous emission of light to some state of lower energy. This fact predicted by S -matrix theory was first noted by Wentzel (1948) in a somewhat different treatment. The connexion between the line width $2b$ for scattering and the coefficient $2b$ for spontaneous emission arises in S -matrix theory from the complementary interpretations of the singularity in the complex energy plane.

As E^0 increases through the value $H_S(\alpha'')$, an additional term appears to be included in the sum for b . This term depends primarily on the factor $|\langle k | V | \mathbf{p} \alpha'' \rangle|^2$, which is zero for photon scattering, since it corresponds to emission of a photon from an atom without a change of state. Thus the matrix element $\langle \mathbf{p} \mathbf{l} \alpha^0 | S | \mathbf{p}^0 \mathbf{l}^0 \alpha^0 \rangle$, calculated in resonance scattering approximation, is an analytic function of E^0 at $H_S(\alpha'')$. As noted above this result is required for consistency with the analytic form of the unitary condition.

The emission coefficient $2b$ can be broken up into separate terms, such as

$$4\pi^2 e^2 \nu^3 \iint \sum_1 |\langle \alpha'' | x_1 | \alpha' \rangle|^2 \sin \omega d\omega d\chi. \quad (7.4)$$

This is the Einstein coefficient for spontaneous emission of a photon of frequency $\nu(\alpha''\alpha')$, with decay of the atom from $|\alpha''\rangle$ to $|\alpha'\rangle$. Knowledge of all the Einstein coefficients is required in order to predict the relative intensities of spectral lines, but the S matrix appears to give only the sum of these coefficients for decay from each level $H_S(\alpha'')$. In order to obtain these relative intensities from the S matrix, it is necessary to consider other matrix elements than the elastic scattering elements. It is convenient to picture a Gibbs ensemble made up of a superposition of individual independent collisions, and the whole representing a stationary scattering process.

The discrete lines in a spectrum are an approximate concept obtained from sharp maxima in a continuous spectrum. To the same approximation the energy of $|\alpha''\rangle$ may be regarded as a sharp value $H_S(\alpha'')$. Then for an energy value such that $H_S(\alpha'') > E^0 > H_S(\alpha')$, an incident photon may excite the atom from $|\alpha^0\rangle$ to $|\alpha''\rangle$. Let x photons of the ensemble do this. The state $|\alpha''\rangle$ will then decay, either by emission of one photon frequency $\nu(\alpha''\alpha^0)$, or by emission of two photons $\nu(\alpha''\alpha')$ and $\nu(\alpha'\alpha^0)$. The latter process takes place by decay first to $|\alpha'\rangle$, and subsequent decay of $|\alpha'\rangle$ to $|\alpha^0\rangle$. Since the ensemble as a whole represents a stationary state, the number y of photons $\nu(\alpha''\alpha')$ leading to formation of $|\alpha'\rangle$ from $|\alpha''\rangle$, must be equal to the number of photons $\nu(\alpha'\alpha^0)$ giving decay from $|\alpha'\rangle$. Further, if z is the number of photons $\nu(\alpha''\alpha^0)$, in order that the probability of $|\alpha''\rangle$ should be constant, $x = y + z$.

Omitting the polarization vector $\mathbf{l}$ for brevity, the S matrix will have the following matrix elements resulting from an initial state $|\mathbf{p}^0\alpha^0\rangle$:

$$\left. \begin{aligned} (a) \quad & \langle \mathbf{p}\alpha^0 | S | \mathbf{p}^0\alpha^0 \rangle, \\ (b) \quad & \langle \mathbf{p}\mathbf{p}(\alpha'\alpha^0)\alpha^0 | S | \mathbf{p}^0\alpha^0 \rangle, \\ (c) \quad & \langle \mathbf{p}\mathbf{p}(\alpha''\alpha^0)\alpha^0 | S | \mathbf{p}^0\alpha^0 \rangle, \\ (d) \quad & \langle \mathbf{p}\mathbf{p}(\alpha''\alpha')\mathbf{p}(\alpha'\alpha^0)\alpha^0 | S | \mathbf{p}^0\alpha^0 \rangle. \end{aligned} \right\} \quad (7.5)$$

The process of decay of $|\alpha''\rangle$ should not be confused with induced emission, which is essentially a non-stationary process, and occurs only when a photon collides with an atom already in an excited state.

Denoting the S matrix elements corresponding to (7.5) (a) to (d), by r_a, r_b, r_c, r_d , respectively, the cross-section for the transition $|\mathbf{p}^0\alpha^0\rangle$ to $|\mathbf{p}\mathbf{p}(\alpha''\alpha^0)\alpha^0\rangle$ will be determined by $|r_c|^2$. Now, the probability of formation of the real intermediate state $|\alpha''\rangle$ is the same for both (c) and (d). Thus the ratio $|r_c|^2 : |r_d|^2$ is determined solely by the relative probabilities of decay of $|\alpha''\rangle$ direct to $|\alpha^0\rangle$, and decay via $|\alpha'\rangle$ to $|\alpha^0\rangle$. In a stationary process each decay to $|\alpha'\rangle$ is associated with a unique decay $|\alpha'\rangle$ to $|\alpha^0\rangle$, since otherwise the probability of $|\alpha'\rangle$ would either increase or decrease. Hence, apart from weighting factors, the ratio $|r_c|^2 : |r_d|^2$ gives the relative intensities of the spectral lines of frequencies $\nu(\alpha''\alpha^0)$ and $\nu(\alpha''\alpha')$. Similarly, all other ratios of intensities may be obtained from the S matrix.

In an exact treatment of photon scattering, one would have to consider matrix elements corresponding to creation of many photons; indeed, for any given energy

an infinite number of photons is theoretically possible. It seems probable that approximate methods will be necessary in S -matrix theory to limit the complexity of the processes considered. This difficulty does not arise, however, for particles of finite mass, whose number is limited by the total energy of the system. But for such particles the difficulty of non-analytic behaviour of matrix elements is encountered. The examples considered in this paper indicate that a new postulate will be required in S -matrix theory, which states how the analytic behaviour changes at threshold values.

I wish to express my thanks to Prof. P. A. M. Dirac and to Prof. W. Heisenberg for valuable criticism and advice, and also to Dr N. Kemmer for helpful discussions.

REFERENCES

- Dirac, P. A. M. 1948 *Principles of quantum mechanics*, 3rd ed. Oxford University Press.
Dyson, F. J. 1949 *Phys. Rev.* **75**, 1736.
Eden, R. J. 1949 *Proc. Roy. Soc. A*, **198**, 540.
Heisenberg, W. 1943 *Z. Phys.* **120**, 513, 673.
Heisenberg, W. 1944 *Z. Phys.* **123**, 93.
Heisenberg, W. 1949 *Two lectures*. Cambridge University Press.
Hu, N. 1948 *Phys. Rev.* **74**, 131.
Hu, N. 1948a *Proc. R. Irish Acad.* **52 A**, 51.
Møller, C. 1946 *K. Danske vidensk. Selsk. Skr., Mat.-Fys. Medd.*, **22**, no. 19.
Touschek, B. 1947 Unpublished.
Wentzel, G. 1948 *Helv. phys. Acta*, **21**, 48.

The organization and work of the Division of Biochemistry and General Nutrition of C.S.I.R.

By H. R. MARSTON,* *Chief of Division*

(Lecture delivered 18 November 1948—Received 16 May 1949)

[Plates 12 to 18]

INTRODUCTION—ORGANIZATION AND ADMINISTRATION

In allotting me the task of relating to you something of the organization and work of the Division of Biochemistry and General Nutrition of the Council for Scientific and Industrial Research, I trust it is not your pleasure to have me dwell on the first part of this title—for a description in measured terms of a substance so tenuous is beyond me. Organization (in its official sense) is not readily discernible to those of us who are gathered together in this laboratory. It is there, but it is never obtrusive.

We are an organ of a body which, although nurtured by Government funds, is free from the vegetative nervous system of the public service administration. We are situated within the precincts of a university, and enjoy the amenities and intellectual discipline of academic life without having to shoulder too great a burden of teaching. Our reference is to conduct research (with a quiet aside that applied science may remain healthy and productive only when thoroughly fertilized with more fundamental studies). Here is a grand experiment in Government administration.

The burden of responsibility entailed is apparent. This, together with a simple faith that practical application will flow freely and naturally from more complete understanding of the underlying phenomena, directs our efforts; and a common curiosity welds us, without loss of individuality, into a team. Our accounts are kept in a way which is acceptable to the Treasury.

The team is composed of individuals whose initial training has been either in physiology or in chemistry—eight senior research men, about twice this number of younger men (the heart of the team), and about four times as many technical assistants (enough to provide all the extra hands and eyes the team may use effectively without being overwhelmed by problems of management). The backbone of the team consists of individuals who have had post-graduate training in British universities and who are imbued with the ideals of which this place is the Mecca.

THE SCOPE OF THE WORK

The meagre fund of exact knowledge of ruminant physiology—which, in light of the importance of sheep and cattle to the well-being of mankind, is rather a depressing heritage—prompted our attention to this general field. Our interest became

* Elected F.R.S. 17 March 1949.

fixed on the nutritional biochemistry of the sheep—a convenient and tractable ruminant—and more especially on factors which influence wool production, for here was a virgin field that might appropriately be cultivated in Australia.

There is no great mystery about the nutritional factors which govern wool production by normal, healthy sheep. The quota of the amino-acids absorbed from the intestine that ultimately serves as a substrate for wool production must escape alike the demands of competing syntheses, and of deamination and utilization as a source of energy—modifying influences which vary with the nutritional state of the animal.

A series of experiments in which there was a precise knowledge of the overall energy transactions, of the nitrogen balance, of the amino-acids in the fodder and in the main synthetic products—wool and flesh, and of the rate of wool production, allowed the influence of various states of nutrition on the efficiency with which amino-acids are utilized for wool production to be assessed with satisfactory accuracy. Appraisal of the nutritional limitations imposed by grazing conditions then became feasible and we were able to proceed intelligently with their correction.

The estimation of the overall efficiency of wool production was greatly simplified by a study of the amino-acid constitution of the vegetable proteins which normally provide the substrate from which wool is elaborated, for this revealed that the protoplasmic proteins, apparently of all orders of higher plants, are of extraordinarily similar composition. The biological value of these proteins for wool production is restricted by their potential cystine content to approximately 35 % under ideal conditions. The theoretical nutritional state in which there would be a complete conversion of the limiting essential amino-acid, however, is never closely approached—even under nutritional conditions in which the energy requirements are met with substances other than protein, and there is a minimum of depletion of the substrate of amino acids by competing syntheses, the efficiency with which the highly evolved Merino sheep converts to wool fleece the amino-acids absorbed from its intestine exceeds 10 % only in exceptional instances. Obviously wool keratin is an expensive biological product when compared with the proteins of meat, milk, eggs, etc.

Usually the concentration of protein in pastures falls very short of that necessary to provide for maximum fleece production, and for this reason the grazing Merino rarely expresses its full hereditary propensity—fine wool, more often than not, is mainly the result of suboptimum nutritional conditions.

The experimental work which is proceeding on the thermodynamics of food utilization, designed primarily to provide a sound basis for maintaining flocks over periods of drought, must be passed over with mere mention. Although energy metabolism was the central theme of physiology at the close of last century, it had been abandoned in a state of no little confusion. About ten years ago we built two calorimeters for the study of energy transactions in the sheep and these tools have since been used continuously to provide an essential basis for our nutritional investigations. Discussion of the work on the intermediary metabolism of the simple fatty acids which, as products of microbial dissimilation of carbohydrates, constitute the main and apparently wasteful source of energy for the ruminant, must be set aside, and so must the studies of the nutrition and activity of the symbiotic microflora of

the paunch, upon which the ruminant largely depends. These projects lack neither scientific interest nor basic nutritional importance, but time is limited and I should like to relate in some detail our experiments with trace elements as these have led in a sort of Pilgrim's Progress (which we like to imagine was a logical sequence) to the solution of agricultural problems of considerable economic significance.

TRACE ELEMENT STUDIES

The geochemistry of the very deficient post-Pliocene aeolian soils

A great part of the soils, along the seaboard of southern Australia are of aeolian origin—their genesis probably dates from a period of extreme aridity in post-Pliocene times.

The coastal fringes are composed essentially of marine shell fragments, wind-borne into gently rolling dunes that cover unconformably and at considerable depth the underlying strata. These merge imperceptibly through consolidated limestone ridges, intersected by brown soils—residua derived apparently from leaching of the older dunes—into a hinterland where light siliceous sands imposed upon solonized subsoils prevail over relatively huge areas.

Indigenous plant associations on these soils

Most of this terrain is in a zone of secure and adequate rainfall. The calcareous littoral supports a very simple plant association in which two calciphile grasses, *Bromus madritensis* and *Lagurus ovatus*, overwhelmingly predominate (plate 12a). The siliceous sands constitute the so-called mallee deserts where a low and stunted heath of dwarfed eucalypts, *Banksia* sp., *Xanthorrhoea* sp., etc., has become inured to the penury of the deficient environment (plate 12b).

Desultory attempts to farm these tracts failed, and all but the marginal areas were abandoned as useless for agriculture.

Coast disease and related maladies

The earliest settlers who ventured to graze their flocks on the natural pastures of the littoral found their enterprise hazardous. Sheep confined to these tracts invariably became anaemic and would waste and die, usually within a year.

This was the first problem that attracted us—it provided the clue which led to the solution of the problem of these refractory lands.

A short preliminary study revealed a syndrome very similar to that of a malady which had been observed to affect sheep depastured on specific areas in most countries, and which had been dubbed with many names. Human experience with this malady had been a long one. Its aetiology was quite obscure.

We proceeded on the assumption that the malady had a nutritional origin. Was it then an uncomplicated deficiency of an essential nutrient element, or was it a deficiency imposed by excessive ingestion of calcium? Experiments soon dispelled the latter idea—and incidentally placed in proper perspective the '*Erdalkali*

Alkalizität' theory which, at that time, tended to obscure our conception of the phosphorus metabolism of ruminants.

The former hypothesis was encouraged when we observed that sheep in an advanced stage of the malady would respond dramatically to dosing with a mixture of soluble salts of the heavy metals (Ni, Co, Mn, Zn, Fe, Cu) which apparently are always present as traces in living matter. As the therapeutic effects were strikingly obvious, cobalt was identified as the active constituent of this mixture within a few months, and a year later we had demonstrated unequivocally that 1 mg. Co/day would prevent the onset of the symptoms which invariably affect sheep depastured on the littoral (figures 1, 2 and plate 13).

The inference that those maladies which are characterized by a similar syndrome might also originate from cobalt deficiency proved correct. A year or so after our first announcement, the New Zealand bush sickness which affects sheep on soils derived from acidic volcanic ejection, the Western Australian enzootic marasmus which affects sheep on certain red loams, the vinkish and daising of the Cheviots, the pine of the raised beaches of the Inner Hebrides, the salt sickness of the coral sands of Florida, and similar maladies elsewhere were all demonstrated, by others, to respond dramatically to cobalt.

But the disabilities suffered by sheep depastured on the deficient littoral were of more complex origin. Another syndrome, previously masked by the more rapidly fatal effects of cobalt deficiency became evident in the animals treated with cobalt. This proved to be the result of copper deficiency.

A further four years of experimental observations conclusively demonstrated that adequate supplements of cobalt and of copper would maintain sheep in normal health and productivity while depastured on this terrain; and collateral chemical evidence left no doubt that the malady was the result of a dietary deficiency of both cobalt and copper. Here, then, where the dual deficiency prevails, appropriate treatment would reveal the uncomplicated syndrome of either deficiency (figure 3).

A unique opportunity was thus provided for intensive study of these deficiencies.

Cobalt deficiency

We found it possible to produce the deficiency syndrome under laboratory conditions by confining sheep to concrete pens, and feeding them on rations consisting of hay produced on the deficient terrain, supplemented with washed gluten and cod-liver oil—a diet which, when further supplemented with cobalt and copper, fulfilled their nutritional requirements. This procedure rendered possible a close study of the malady (figure 4).

A sheep exhibits characteristic symptoms when cobalt deficient (plate 13). Its normal alert demeanour changes to one of dull and rheumy-eyed listlessness; its buccal and conjunctival mucosae blanch; its skin pales from the normal ruddy pink to a dull greenish hue, becomes less flexible and in the advanced stages of the malady sufficiently fragile to break on parting the wool. A steady decline in body weight reflects a failing appetite which progresses until the animal dies of inanition. Autopsy reveals little that is distinguishable from hunger oedema—there was little evidence which might provide a clue to the physiological mechanisms involved.

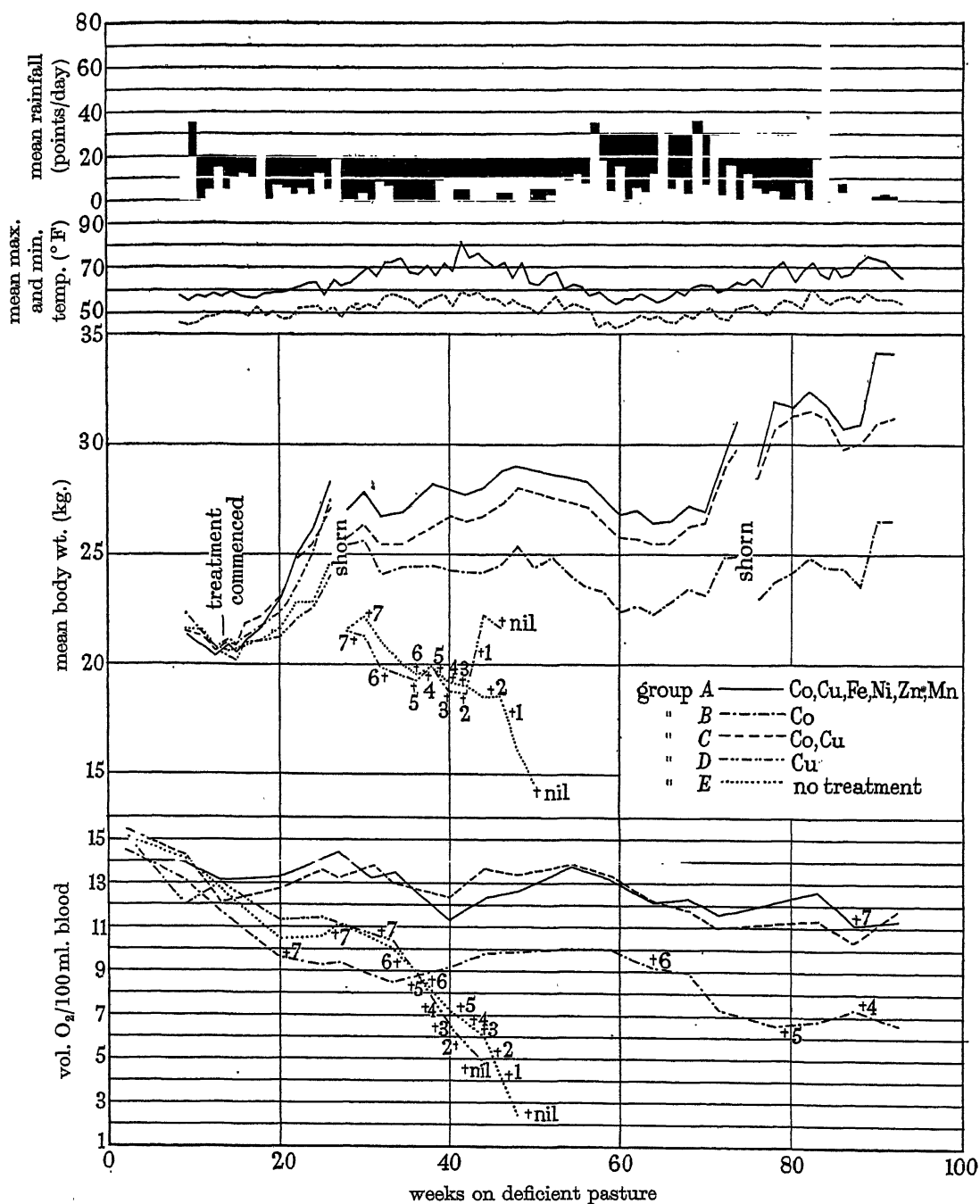


FIGURE 1. The data are from an experiment in which an even flock of ewes (*aet.* 9 months at the beginning of the experiment) was divided into 5 groups each of 8 animals and treated, while depastured on the deficient littoral, by dosing thrice weekly with the elements indicated. The effects of additional cobalt are obvious. Those without it all succumbed within a year. The effects of copper deficiency previously masked by the more rapid effects of cobalt deficiency become apparent in group B. Other elements superimposed on cobalt and copper exerted no significant benefit. (From experiments of Marston, Lee & McDonald 1936.)

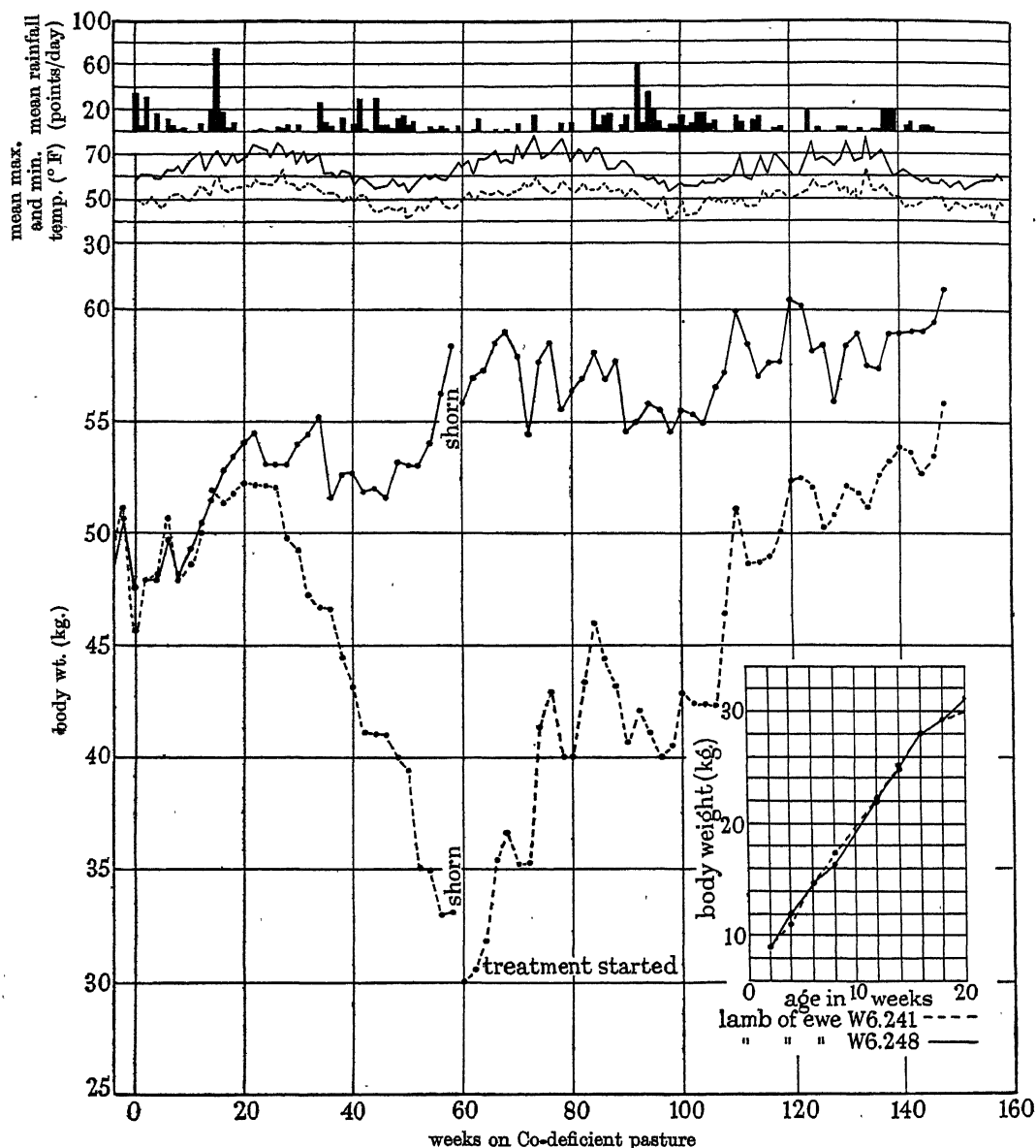


FIGURE 2. The data refer to one of a number of pairs of mature twin Merino ewes which were transferred to the deficient littoral in 1936. Ewe W 6.248 was treated with cobalt throughout the experimental period of 3 years. Ewe W 6.241 was untreated during the first 55 weeks after which a supplement of cobalt sulphate equivalent to 1 mg. Co/day was administered. Both ewes received 10 mg. Cu/day. The supplement of cobalt prevented loss of weight and the appearance of deficiency symptoms in ewe W 6.248. The effect of cobalt administered to ewe W 6.241 while in the terminal stage of the deficiency syndrome (see plate 13) illustrates the typical therapeutic action. There are no striking permanent effects after recovery from an extreme state of cobalt deficiency. The growth rates of the lambs (inset) produced subsequent to mating these ewes at 110 weeks are practically identical. Both received 1 mg. Co/day. —•—•— ewe W 6.248 (treated), - - - • - - ewe W 6.241 (untreated).

The anaemia which is a feature of the malady presents a blood picture suggestive of an aplastic condition of the bone marrow (plate 13*b*). In the early stages, a reduction in the number of cells rather than changes in their mean size or in their mean haemoglobin content contributes to the low oxygen-carrying capacity. Later, as the state of anaemia worsens, the picture becomes one of a profound large-cell anaemia, somewhat reminiscent of, but quite distinct from that of Addison's pernicious anaemia of man.

All attempts to produce cobalt deficiency in small laboratory animals (rats, mice, rabbits, etc.) have failed—most of the experiments reported in the literature are

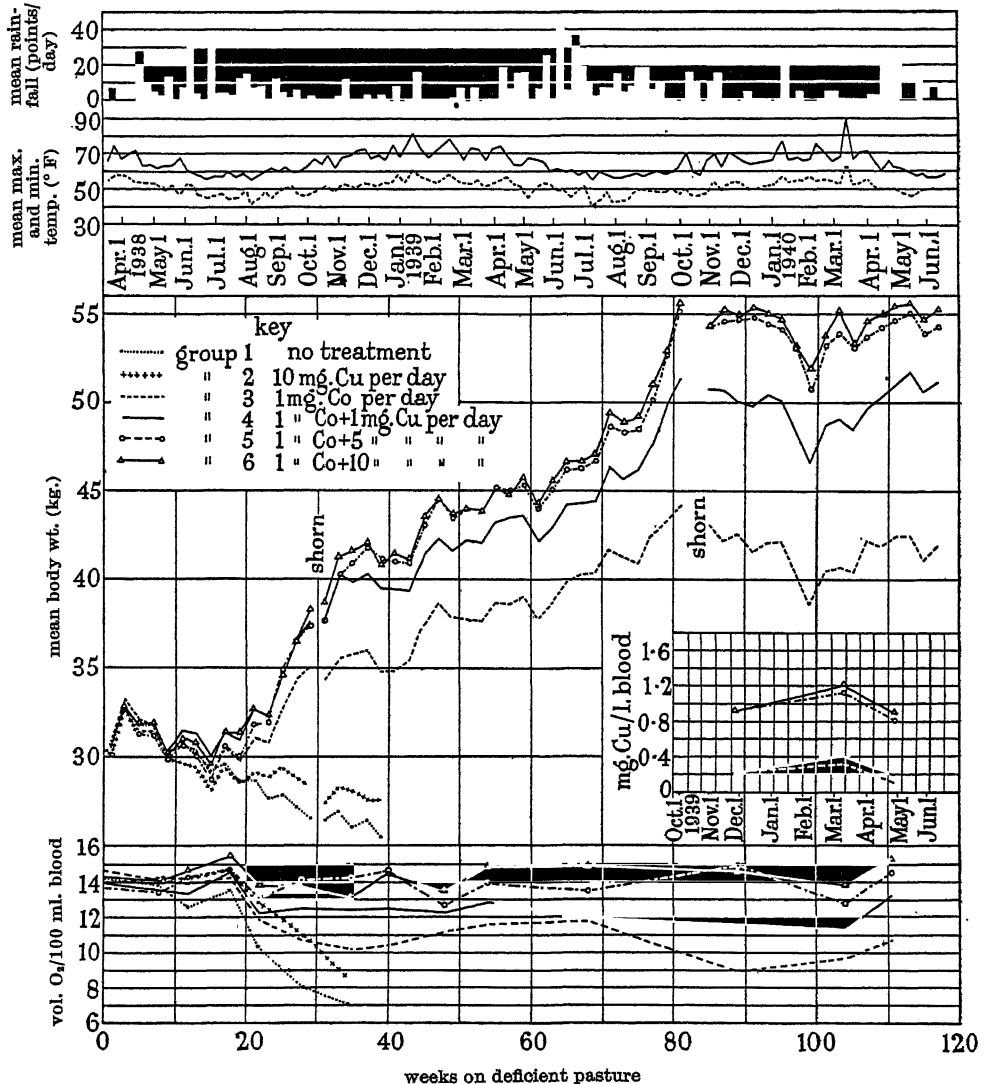


FIGURE 3. The data are from an experiment designed to estimate the quantity of copper necessary to fulfil the requirements of sheep depastured on the deficient terrain. This flock was divided into six groups each of 10 animals and treated as indicated. (From experiments of Marston, Lee & McDonald 1938.)

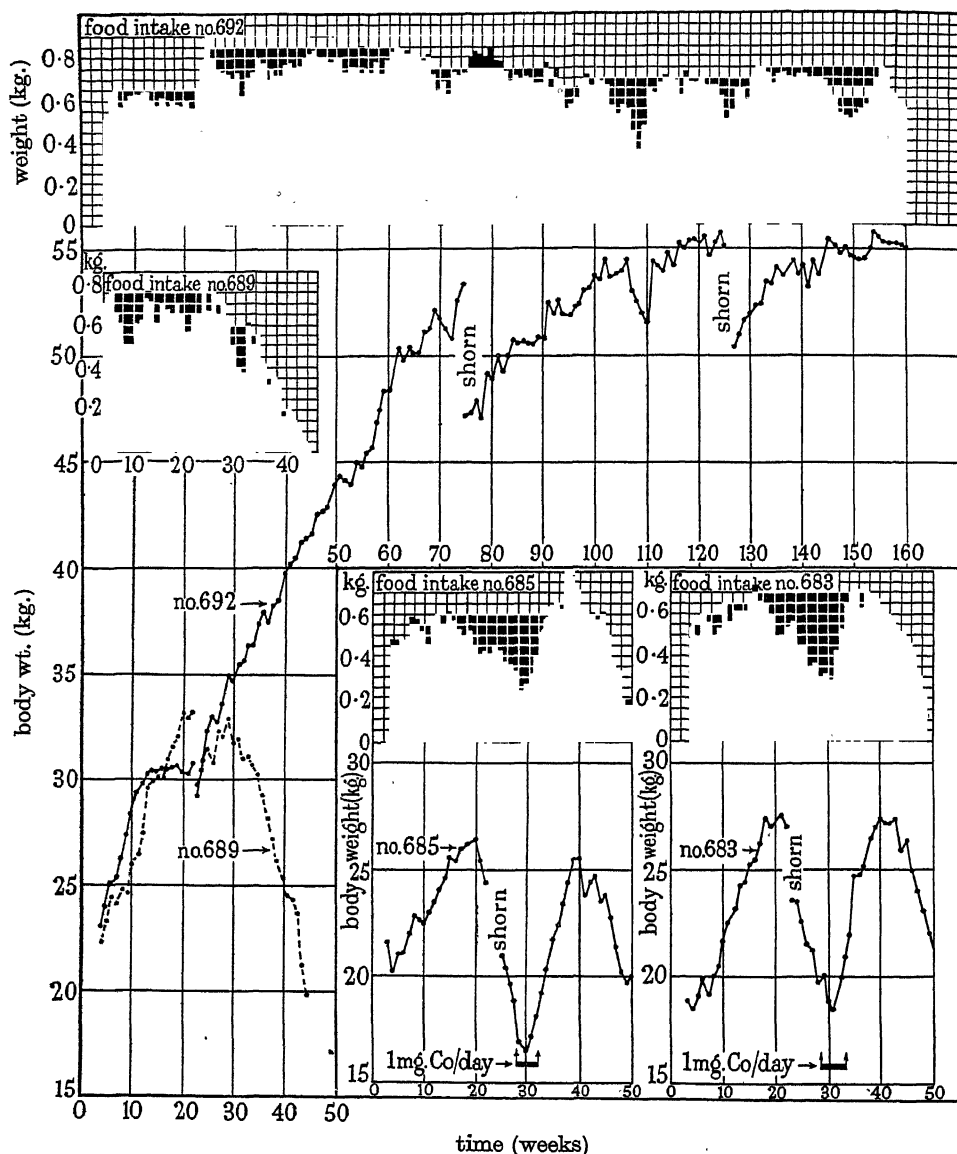


FIGURE 4. Rye hay grown on the deficient littoral supplies adequate nourishment for sheep if provided *ad lib.* and supplemented each day with 100 g. washed gluten, 4 ml. cod-liver oil and 1 mg. Co and 10 mg. Co. Without additional cobalt this diet is inadequate—sheep fed on it develop progressive symptoms of cobalt deficiency which terminate fatally. The data illustrate the effect of the cobalt supplement on sheep that have been pen-fed on this diet. The growth rate of no. 692 which received 1 mg. Co/day from the beginning of the experiment is normal. No. 689, to which no extra cobalt was supplied, developed well-defined symptoms of cobalt deficiency, lost weight and died 44 weeks after being confined to this diet, the course of the syndrome being identical with that observed in the field.

The insets show the effects which supervened when 1 mg. Co/day was provided for 4 weeks to sheep which had developed the deficiency syndrome under these conditions. Note effects on food intake which is represented as mean daily intake of food/week. The rye hay contained $0.02 \mu\text{g. Co/g. dry wt.}$; the gluten $0.02 \mu\text{g. Co/g. dry wt.}$. The water provided for the animals was distilled. (From unpublished experiments of Marston & Peirce 1938-9.)

unconvincing, however, as they have been of short duration and the chemical claims have been of doubtful significance; but we have now the third filial generation of rats that have been sustained, with elaborate precautions and meticulous chemical control, on a diet containing < 0.03 parts of cobalt per million dry weight, upon which they remain normal and healthy. Sheep would certainly succumb to the deficiency within a year if confined to fodder containing this concentration of cobalt.

The amount of cobalt necessary to overcome the fatal consequences of the deficiency is certainly very small. The equivalent of 1 mg. Co/day which was arbitrarily adopted in the first series of experiments proved on further study to be in large excess of the actual needs of sheep depastured on the deficient littoral where the intake in the fodder varies between 20 and 30 μg . Co/day—the equivalent of 0.1 mg. Co/day was sufficient to prevent entirely the appearance of any untoward symptoms, *provided that it was administered frequently*. In these experiments the positive controls were dosed thrice weekly; correspondingly larger doses given at intervals of two weeks failed to prevent the onset of deficiency symptoms, and equivalent doses administered at intervals of 6 weeks merely delayed their fatal termination.

The fact that sheep and cattle are apparently the only animals which suffer the consequences of cobalt deficiency while grazing on the deficient terrain—marsupials abound, the rabbit population if unchecked assumes plague proportions, and horses do exceptionally well there—suggests that the ruminant is either extraordinarily wasteful with cobalt or is unique in its greater demand for this element. Our observations support the latter hypothesis.

In our first experiments with sheep transported from normal pastures to the deficient littoral we had observed the deficiency syndrome develop to its fatal conclusion in some whose livers, when analyzed subsequent to death, had retained as much cobalt as is found in the livers of animals grazed elsewhere on normal pastures. Later, we established depots in the sheep by introducing relatively massive quantities of slightly soluble cobalt salts under the skin, and observed no benefit from this procedure. On investigating the reason for this, we obtained unequivocal evidence that if cobalt is to be effective it must be ingested—when injected into the blood stream it is valueless to the animal (figure 5) although the cobalt concentration in the organs may by this means be increased at least ten-fold above that found in normal healthy sheep. Clearly then, cobalt exerts its action primarily either in the lumen of the alimentary canal or during its passage through the wall. The site of this activity appears to be at a level above the duodenum, for study of the fate of Co^{60} injected intravenously revealed that a considerable proportion of cobalt introduced in this way is voided in the faeces; it is not secreted in the saliva but finds its way via the bile into the intestine from which it is not reabsorbed.

An important function of cobalt in the ruminant might thus concern the symbiotic flora of the paunch upon which the nutrition of the ruminant depends. If the demands of these micro-organisms are large compared with the requirements of the animal itself, then the strikingly dissimilar behaviour between ruminants and other herbivora might be explained.

Attention has now been fixed on the essential nature of cobalt for the nutrition of animals other than ruminants, by the announcement a few months ago of the isolation of Minot's anti-pernicious anaemia factor from liver, and of the discovery, both in the United Kingdom and in the United States, that this bright red, crystalline material is an association complex of cobalt.

These findings already suggest that the quantity of cobalt required for the normal physiological function of the bone marrow is minute, as $<10\mu\text{g.}$ of this factor (equivalent to $<0.4\mu\text{g. Co}$) introduced parenterally has been reported to correct for at least 7 days the dyscrasia of human pernicious anaemia.

Deficient pastures, which are never completely devoid of cobalt (they contain between 0.02 and $0.03\mu\text{g. Co/g. dry wt.}$), may provide to all grazing animals sufficient cobalt for the elaboration of this factor—possibly by interaction with Castle's extrinsic factor in the acid gastric contents—but leave unfulfilled, in the ruminant, the requirements of other physiological mechanisms (those of the symbiotic microflora for example) which demand greater quantities of cobalt for their normal activities.

The profound anaemia that supervenes on cobalt deficiency in the ruminant has certain features in common with other macrocytic anaemias of nutritional origin. But its origin is apparently not due primarily to a deficiency of the A.P.A. factor. Our current experiments suggest that it does not respond either to massive amounts of liver extracts administered parenterally or to similar therapy with the pure A.P.A. factor. Nor will it respond to dosing *per os* with pteroyl-glutamic acid. In so far as our experience extends, cobalt administered *per os* (as a soluble ionizable salt) is the only effective agent which will prevent the onset of the deficiency syndrome in sheep depastured on the deficient tracts, or will cure the symptoms when they have developed.

The anaemia which is a feature of the cobalt-deficiency syndrome in the ruminant is apparently the result of a breakdown of a link in the chain of metabolic events responsible for the maturation of the red cell, other than the one that fails in human pernicious anaemia.

In contrast to the dramatic change in the general health and activity of a deficient sheep which supervenes immediately on treatment with cobalt, the recession of the anaemia is slow, and so the low oxygen-carrying capacity of the blood *per se* is not the prime cause of the physiological disability but a symptom of a more general metabolic defect. *It appears that we may seek with confidence a less restricted physiological role for cobalt than the specific part it plays in the maturation of the red cells.*

And that is the present state of our knowledge of the function of cobalt in the ruminant. Our curiosity is now directed to the possible significance of cobalt in the nutrition of the associative growth of micro-organisms which normally inhabit the rumen. The A.P.A. factor is known to be necessary for the growth of the Dorner strain of *Lactobacillus lactis* and this implies that cobalt may be an essential adjuvant for other bacteria—the quantity required, $1.3 \times 10^{-5}\mu\text{g.}$ of the crystalline factor (equivalent to $5 \times 10^{-7}\mu\text{g. Co}$) per ml. of medium for half maximum growth, suggests that the A.P.A. factor is among the most potent of the compounds essential for the proliferation of micro-organisms.

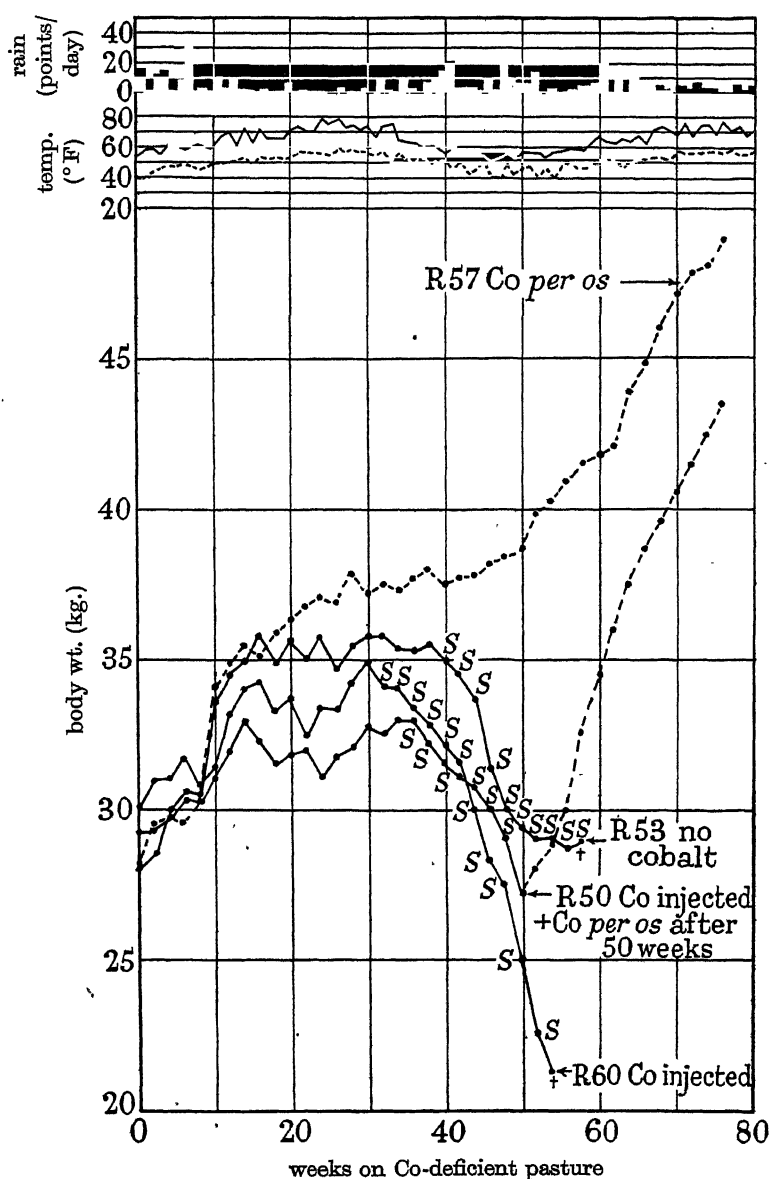


FIGURE 5. The individuals of this group of 4 Merino ewes (*aet.* approx. 9 months at beginning of the experiment) were grazed together on cobalt-deficient pastures and treated as follows: R. 57, with the equivalent of 1 mg. Co/day administered once a week *per os*; R. 60, and R. 50, with 1 mg. Co/day injected once a week into the jugular vein; R. 53 received no supplement of cobalt. All received adequate copper. After 50 weeks, when R. 50 was in the terminal stages of the deficiency syndrome, the equivalent of 1 mg. Co/day was superimposed on the injections. S denotes recognizable deficiency symptoms. The behaviour was similar to that of the animals in other groups. The cobalt concentration of the pastures varied between 0.02 and 0.03 $\mu\text{g. Co/g. dry wt.}$ (From experiments of Marston & Lee 1943-44.)

Copper deficiency

It is possible to be more concise and somewhat more final in the discussion of our experience with copper deficiency—for reasons which may later become clear, problems associated with the physiological role of copper were pursued with more vigour and they have yielded more readily than those that concern the function of cobalt.

When the cobalt requirements of sheep depastured on the deficient littoral are fulfilled, a syndrome slowly develops which has proved to be the result of a serious copper deficiency; additional copper leads rapidly to its regression, and will maintain the animals in normal health and productivity.

Sheep are unable to establish a positive copper balance on these deficient pastures which contain, according to the stage of growth, between 2 and 4 $\mu\text{g. Cu/g. dry wt.}$ Their reserves of copper are depleted and the reduced status is reflected by a steady fall in the blood-copper to levels below one-fifth of the normal concentration of approximately 1 $\mu\text{g./ml.}$ They sicken and exhibit well-defined symptoms. By appropriate treatment with supplements of copper insufficient to maintain the normal physiological requirements, we were able to produce an ordered series of deficiency states that ranged from the frank syndrome in the untreated animals to an incipient state of deficiency in which overall efficiency was impaired without the appearance of specific symptoms (figure 6).

It was our experience with this series of experiments which led us later to recognize the aetiology of a number of ill-defined maladies of young grazing stock, enzootic over extensive areas distinct from the very deficient terrain upon which these observations were made. And they provided the second, and this time, much more definite clue to the understanding of the factors which had rendered agriculture impossible on the vast areas of the so-called mallee deserts.

The observations showed clearly that when the copper status of the sheep falls below a well-defined level, the first metabolic process to become seriously impaired is the chemical mechanism concerned with keratinization. Further depletion results in a disordered iron metabolism which becomes manifest in a hypochromic anaemia and the deposition of haemosiderin in the tissues. A very low copper status in the ewe during pregnancy leads invariably to nervous sequelae in her lamb.

Nervous sequelae of copper deficiency

The results of this nervous disorder have been known to shepherds for many centuries in the British Isles and elsewhere.

The symptoms range, according to the severity of the lesion, from complete paralysis at birth to a bland ataxia which may appear at any age up to 6 months or more. The incoordination of the hind limbs in the milder cases presents a clinical picture which suggests failure of the upper motor neurones; and histological examination of the cord reveals extensive demyelination in the ventro-lateral column always involving the direct and crossed pyramidal tracts which are not as highly differentiated in the sheep as in the primates.

The invariable nature and restricted distribution of these lesions raises many

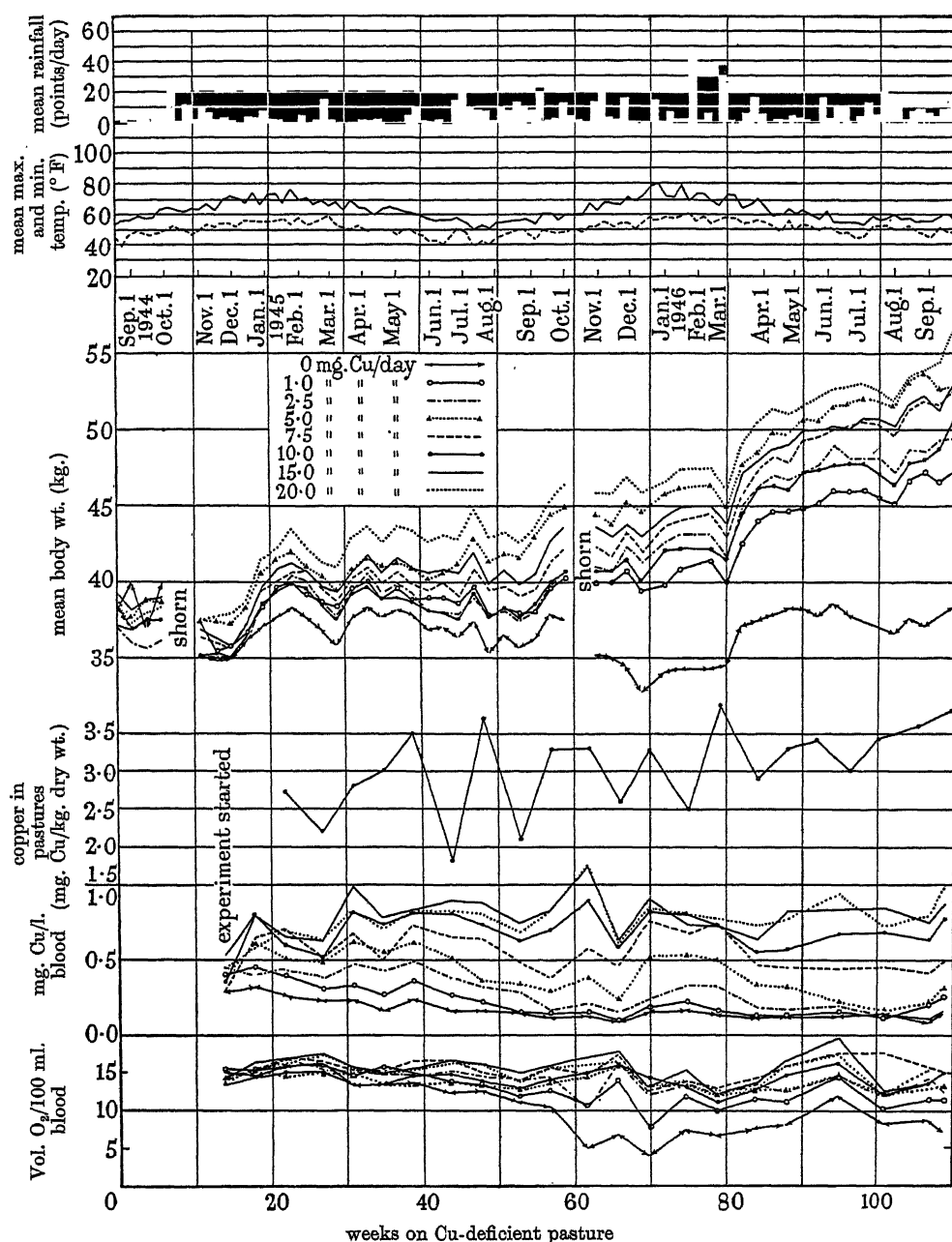


FIGURE 6. The data are from an experiment designed primarily to study the influence of various degrees of copper deficiency on the quantity and nature of the wool produced by Merino sheep. This flock was selected from a large number of young Merino ewes to be of even wool type. It was divided into 8 groups each of 8 animals and grazed on seriously copper-deficient pastures. All received cobalt. The influence of the supplements of copper on the oxygen-carrying capacity and on the copper concentration of the bloods is shown. The effects on the wool fleece are indicated in figure 8 and plate 15. (From experiments of Marston & Lee 1945-6.)

intriguing questions, but it is the effects of copper deficiency on the process of keratinization that I should like to relate in some detail to you.

Copper deficiency and keratinization

Copper deficiency reveals itself first as a marked deterioration in the quality of the fleece (plate 14a). A gradual series of morphological changes reflects in the wool staple the progressive decrease in the efficiency of chemical mechanisms in the follicles which require copper if they are to function normally. As the copper reserves are depleted the crimp becomes progressively less distinct until the fibres emerge entirely devoid of character.

The capacity of the follicles to impart crimp returns immediately on the resumption of a normal copper status and the staple that grows subsequently resumes boldly the character of the particular wool type of the individual (plate 14b).

There is thus little doubt that the changes in structure reflected by this lesion in the fleece are due to the breakdown of a catalytic process in which copper is primarily involved. The follicles of a black sheep seriously depleted of copper, lose their ability to impart crimp and their capacity to produce pigment. The role assumed by copper in effecting oxygen transfer is beautifully illustrated by the sudden re-appearance of both melanin and crimp when the normal copper concentration in the follicles is reinstated (plate 14c).

Consideration of the cytology of the follicle leads to the site of this failure. As the proliferating cells at the base of the follicle mature, their expanding protoplasm impels them to present a minimum surface and fill with no interstices the restricted space they occupy so that each cell becomes faceted and approximates to the shape of Kelvin's minimum tetrakaidecahedron. In this state they are forced towards the conical constriction where they are gradually compressed to about one-eighth of their original cross-section and so become greatly elongated into spindle-shaped cortical cells which, when keratinized, comprise the wool fibre. During their extrusion through the constriction the globular protein of these cells granulates and proceeds through microfibrillae of fibrous protein (Vörner's trichohyalin of classical histology) to hard keratin.

A specific histochemical reaction with alkaline nitroprusside which forms a deep purple complex with thiol ($-SH$) groups reveals the process which is impaired by copper deficiency. When this reagent is applied to a normal fibre freshly plucked so as to retain its root, the basal cells are stained very lightly if at all. At the level where the intracellular structures begin to granulate and form into microfibrillae, the staining is intense and from there, in a normal fibre, it extends for approximately 100μ to cease quite abruptly in the fully keratinized portion. When applied similarly to a wool fibre from a sheep with a low copper status the intense reaction begins at the same level but it extends for 1000μ or more—practically the whole of the fibre embedded in the skin reacts strongly, the staining of the upper portion becoming progressively less intense as it approaches the skin surface (figure 7). On dosing the deficient animal with copper this greatly increased zone of free thiol groups retracts to its normal position within a few hours.

The evidence leaves little doubt that copper is primarily responsible for catalyzing the oxidative closure of the thiol residues of the prekeratin fibrous protein to the disulphide linkages of keratin. The rate at which this reaction proceeds is a measure of the copper status of the sheep. In the normal integument the oxidation of the reduced fibrous protein is complete in 8 to 12 hr.—the fibre is set deep down in the follicle: in the integument of a seriously copper-deficient animal this oxidative change takes 3 days or more to proceed to completion and so the fibre retains its plastic condition until it emerges at the skin surface.

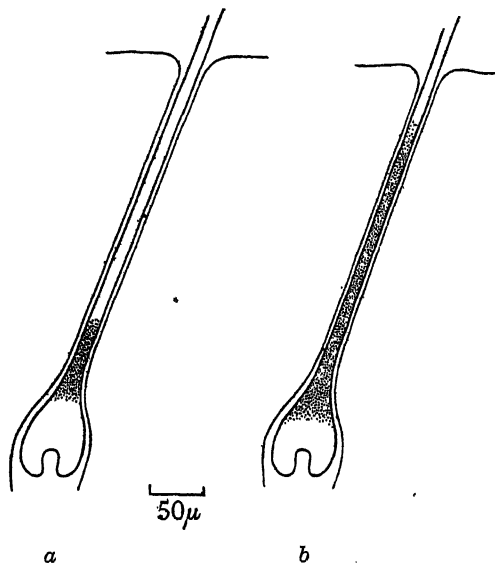


FIGURE 7. Wool follicles (diagrammatic) showing the approximate distribution of free $-SH$ groups in the wool fibre during keratinization as indicated by the nitroprusside reaction applied to wool from *a* normal and *b* copper-deficient sheep.

The preferred molecular orientation in the wool fibre originates during the passage of the cells through the constriction at the base of the follicle—the normal fibre becomes anisotropic during its extrusion through this constriction and the molecular asymmetry remains. The long-chain fibrous protein molecules are aligned while still in a reduced state and fixed in this position by copper-catalyzed oxidative closure of the cysteine residues of adjacent molecules to disulphide bonds which bind the keratin fibrillae together. In the copper-deficient sheep the comparatively slow oxidation of the fibre provides ample opportunity for subsequent disorientation before internal bonding is completed.

The aberrant physical properties of copper-deficient wool—its reduced tensile strength, its lack of elasticity and abnormal stress strain cycle, its tendency to become permanently set, its reduced affinity for dyes, etc.—with little doubt arise from the abnormal internal bonding and less orderly structure within the fibres. There is no gross change in its amino-acid constitution.

When the copper status of a sheep falls below normal, the efficiency of intermediary metabolism is impaired; appetite is lessened and so the rate of wool growth

is diminished. We employed methods analogous to a chemical titration to estimate the extent of this. Supplements of copper ranging in increments from nil to amounts that were more than sufficient to meet the nutritional requirements were administered to a series of carefully matched groups of sheep depastured as a flock on the frankly deficient terrain (figure 6). The mean weight of clean-scoured wool produced was found to vary exponentially with the amount of copper provided in the supplement. In the nutritional environment of the littoral dunes, the asymptote of maximum response was reached when between 7.5 and 10 mg. Cu/day was supplied. Equilibrium was established at this level of intake and greater amounts of copper brought about no further improvement either in the quality or in the quantity of wool produced (plate 15 and figure 8).

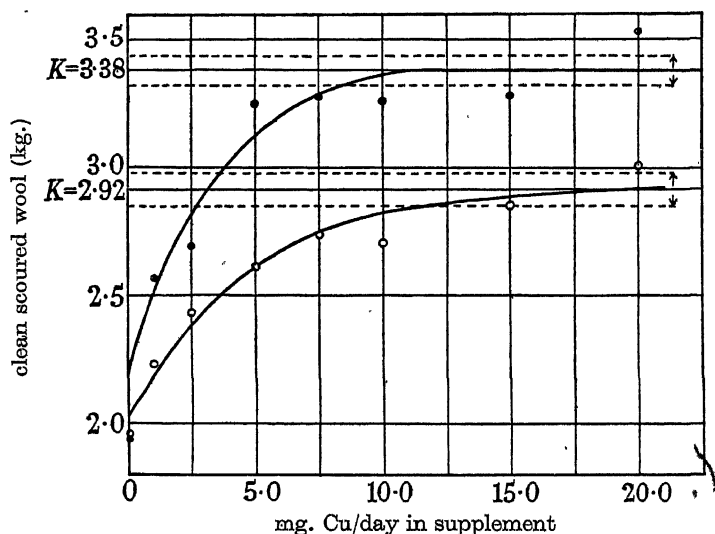


FIGURE 8. The influence of a graded series of copper deficiency (see figure 3 and plate 15) on the annual production of clean scoured wool during the first 2 years of the experiment is shown (○ 1945 clip, • 1946 clip). The observed means for each group are plotted in relation to curves derived by the method of least squares to specify that the increase in wool production per unit of copper supplement is proportional to the deficit below the maximum wool production K . The asymptote K is indicated with its s.d. plotted on either side. In each year the maximum response was achieved when the supplement provided between 7.5 and 10 mg. Cu/day. The pastures were more nutritious in 1946 than in 1945.

Adequate copper more than doubled the monetary value of the fleeces from sheep on this very deficient terrain and the fleeces from the deficient groups ranged in value according to the degree of deficiency imposed. The reduced value of the fleece from a copper-deficient sheep is not due entirely to its lighter weight. The lack of crimp is recognized by wool-buying appraisers to indicate properties that have been found undesirable in technological processing and so copper-deficient wool is heavily discounted.

These titration experiments allowed the consequences of all degrees of copper deficiency to be expressed in language readily appreciated by pastoralists and general adoption of preventative measures followed.



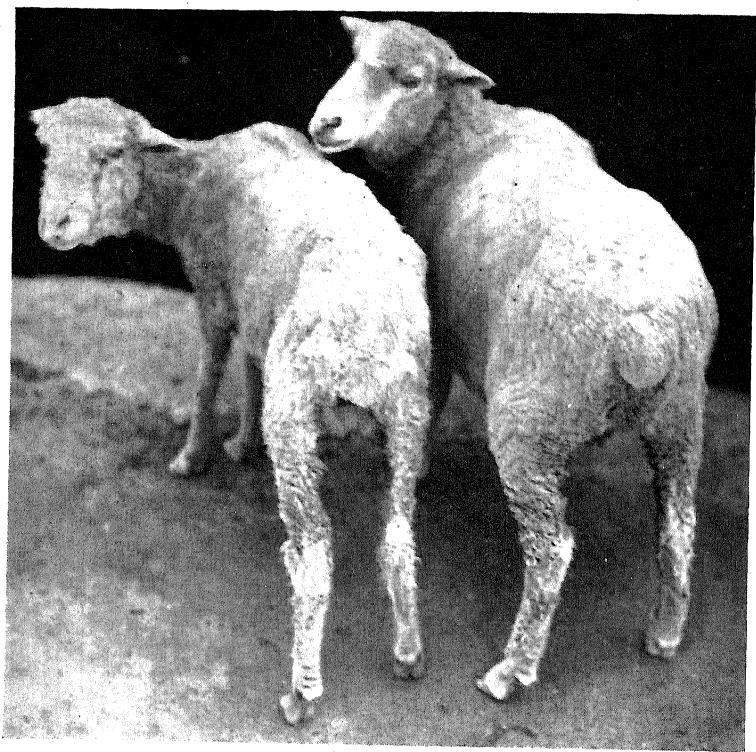
a Typical calcareous dunes



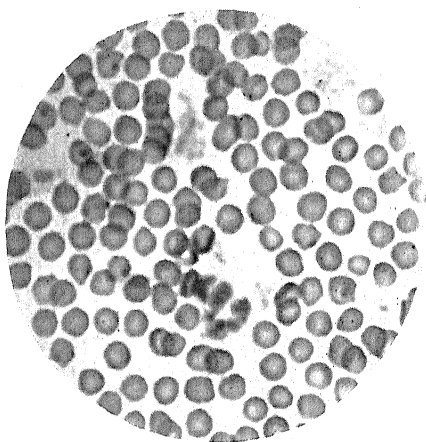
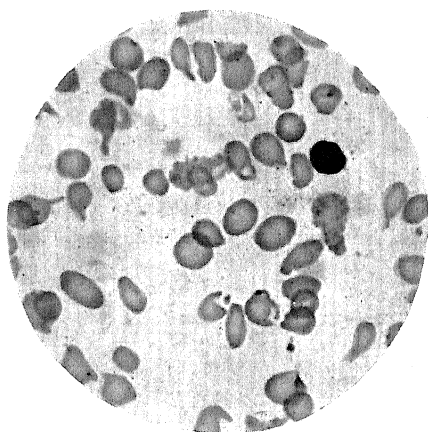
b Typical mallee heath

Areas of aeolean soils in Southern Australia where trace element deficiencies prevail.

(Facing p. 288)

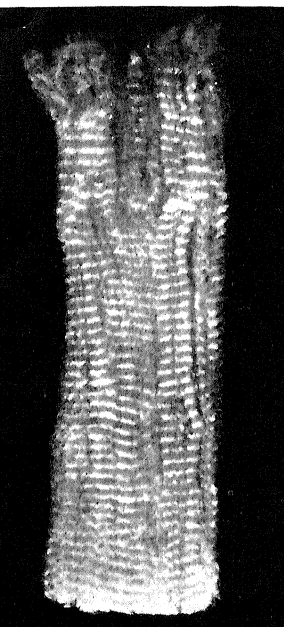


a

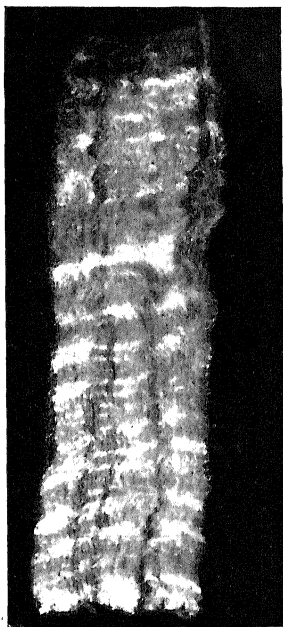


b

The effect of cobalt deficiency on Merino sheep.



1941

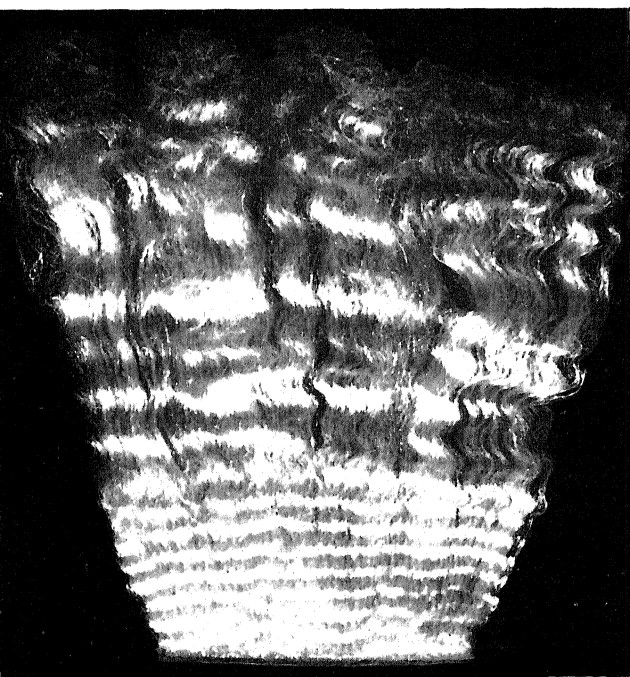


1942

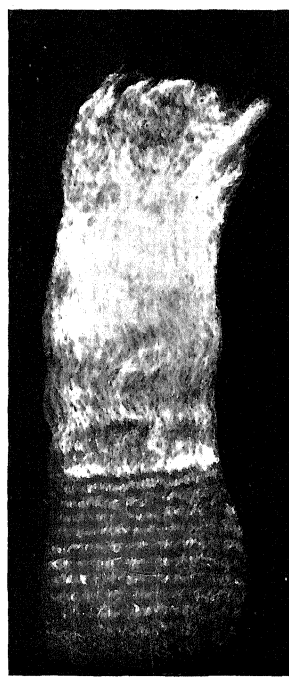


1943

a

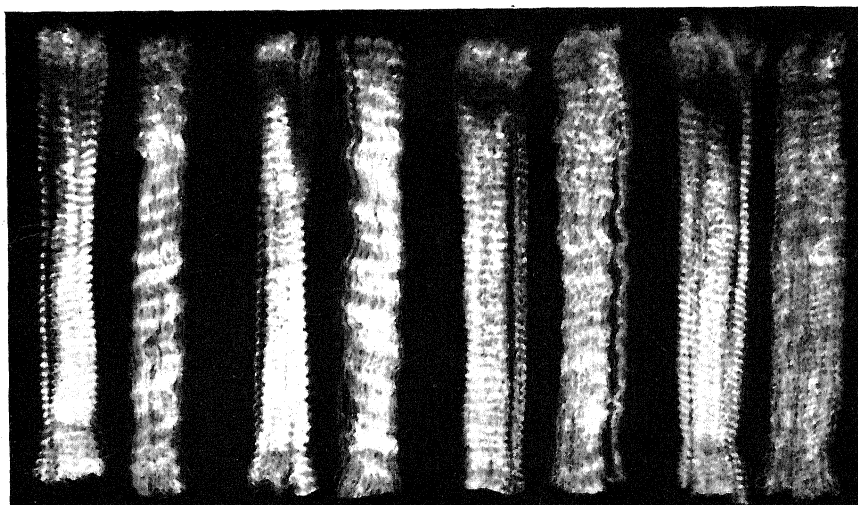


b

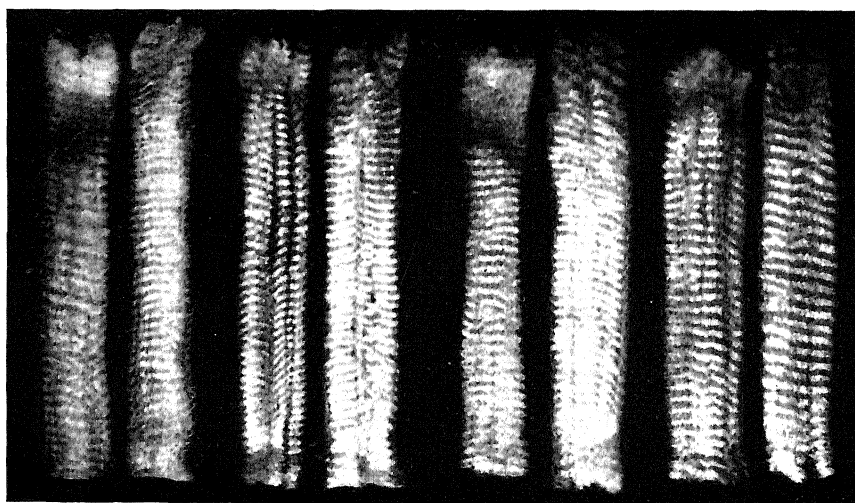


c

The effect of copper deficiency on the wool grown by Merino sheep.



1944 1945 1944 1945 1944 1945 1944 1945
 no Cu 1 mg. Cu/day 2.5 mg. Cu/day 5.0 mg. Cu/day

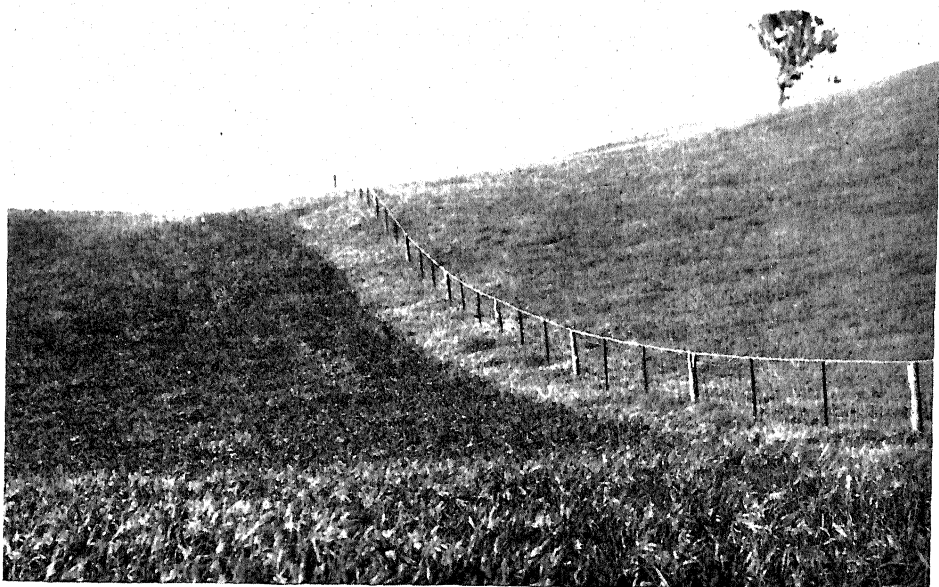


1944 1945 1944 1945 1944 1945 1944 1945
 7.5 mg. Cu/day 10 mg. Cu/day 15 mg. Cu/day 20 mg. Cu/day

The effect of different degrees of copper deficiency on the character of Merino wool.



a



b

Resistance of certain plant species to soil conditions which impose copper deficiency on other species.



a

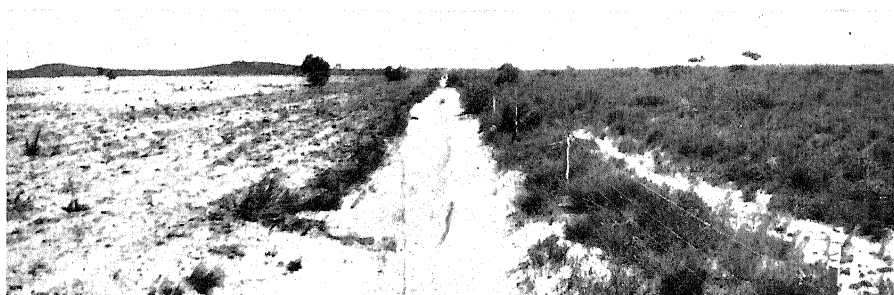


b



c

The establishment of lucerne on the calcareous littoral.



a



b



c

The development of pastures on the mallee heath soils of the hinterland
S.E. South Australia.

Delineation of deficient areas

In the preliminary investigations, geochemistry with its full appreciation of the petrological aspects of soil origin provided valuable guidance, and in virgin country ecological studies have been helpful in delineating deficient areas within any one climatic zone.

On the very deficient tracts where only that vegetation inured by selection to the inadequate nutritional environment is able to persist, the small quantities of micro-nutrients which become concentrated in the fodder plants are grossly insufficient to fulfil the requirements of grazing ruminants.

On less seriously deficient terrain, however, the chemical composition of the available pastures is subject to a great variety of environmental influences: those of climate—the seasonal effects of light, rainfall, temperature, etc.; those of the soils—the nature and interaction of the minerals of the soil proper and of the parent rocks, the pH, the structure, the microflora, etc.; and those of the vegetation itself—the species, variety, strain of the plants, their root systems, their stage of growth, the mineral interactions within them, to mention but a few, are all reflected in the composition of the fodder. And so it is not feasible, under these conditions, to assess the status of micro-nutrient elements without some means of integrating this complex and labile system of variables. Indicator plants do this, and if appropriate species are selected, specific lesions are exhibited. The grazing animal further resolves these factors and their influence may be measured in terms of animal production.

We have exploited specific symptoms such as the copper-deficient lesion in the wool for mapping the areas that are seriously involved and for revealing incipient deficiencies of copper on terrain where hitherto such limitations had never been suspected.

The extent of the areas where less obvious deficiencies are prevalent is very much greater than that of the frankly deficient terrain, and the consequences are of little less economic significance. The recognition and correction of deficiencies on these great tracts of marginal lands followed experiments with indicator plants and led finally to the development of pastures on the very deficient aeolian soils.

Solution of the problems of the aeolian soils

Development of pastures on the littoral

From the beginning the factors which limited so strikingly the development of vegetation on these huge tracts intrigued us.

In our preliminary experiments with animals grazing on the littoral we attempted to establish a sown pasture of a high-yielding perennial grass and a legume suited to this climatic zone by applying liberal dressings of phosphate, nitrogen and potash. The manurial treatment served only to stimulate the grasses already prevalent—the sown species failed completely (plate 16b).

Subsequently, when we realized that copper deficiency in the fodder limited the welfare of animals depastured there, we were led to investigate the effects of increasing the level of available copper in the soils. It was a series of experiments with oats

as an indicator plant that proved the low fertility of these calcareous sands to be due primarily to a low copper status.

The experience of early settlers who had attempted without success to farm these deficient tracts suggested that rye would grow there but oats, wheat and barley would fail. This was substantiated. Oats proved particularly susceptible, for when grown in this environment the oat plant displayed flagrant lesions and failed to head (plate 16a). Plot experiments of factorial design revealed that a dressing of 7 lb. $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ /acre applied at sowing prevented these deficiency lesions and a crop of 23 bushels/acre resulted—nitrogen was then the limiting factor for when nitrogenous manures were superimposed the yield was more than doubled. An identical dressing of copper applied a year before sowing resulted, under the same conditions, in a crop of 62 bushels/acre. Here is a probable effect of copper on nitrification within the soils.

In the absence of additional copper all legumes failed completely, but with it lucerne (plate 17), black medic, barrel medic and other pasture legumes developed with surprising vigour, and under these conditions pasture grasses grew well in association.

The effects of the copper dressings proved lasting—splendid high-yielding mixed pastures established in this way several years ago have persisted without further additions of copper—the yield tending to increase rather than diminish with time.

The factors which had rendered agriculture impossible on about 2000 square miles of littoral were thus revealed and application followed.

Development of pastures on the siliceous hinterland

Although experimental flocks confined to the great hinterland of siliceous sands provided evidence of serious copper deficiency in the sparse herbage on which they subsisted, our experiments there with oats as an indicator proved zinc and not copper to be the first limiting factor.

When the phosphate level of these virgin soils was raised, oats grew and yielded a poor crop of less than 10 bushels/acre. No apparent lesion of copper deficiency appeared and dressings of copper produced no significant response in the yield. But very striking symptoms of disordered metabolism became evident—the older leaves assumed a deep purple colour. A dressing of 7 lb. $\text{ZnSO}_4 \cdot 6\text{H}_2\text{O}$ /acre prevented the accumulation of anthocyanin and doubled the yield before the low nitrogen status of these soils became the limiting factor. Lucerne responded to copper but not to zinc—here again a species difference. Subterranean clover failed completely to develop without additional zinc, grew luxuriantly with it, but in the absence of a dressing of copper failed to set viable seed and so did not persist.

Further series of experiments proved unequivocally that splendid, high-yielding permanent pastures could be established by sowing the perennial grass, *Phalaris tuberosa*, and the annual, subterranean clover, with a dressing of superphosphate enriched with 7 lb. each of $\text{ZnSO}_4 \cdot 6\text{H}_2\text{O}$ and $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ /acre (plate 18), and that—an important economic consideration this—these pastures could be established under a cover crop of wheat or oats, the value of the grain from which would more than discharge all costs of the sowing.

The main factors which have limited the development of vegetation on these huge areas of hungry soils have been revealed, the course that agricultural development should take is clear, and already application is proceeding apace, with dramatic results. About 3000 square miles of sandy heath is involved.

CONCLUDING REMARKS

Although application of these findings met with spectacular success and the implications are of considerable economic importance, there is no need here to confess that application was not the immediate goal for those of us who carried out the experiments. Our delight has been more with the phenomena that have been revealed, with the problems that these have clarified, and with the pattern as it emerged. Each step allowed a glimpse of physiological mechanisms in which minute traces of heavy metals play an essential part. The urge to add further to this picture will continue to be our main stimulus: I have no doubt that pursuit of knowledge along these lines will pay economic dividends.

The realization that research in the applied field may miss the substance for the shadow if it becomes too practical, was the central thought in the minds of the small group who launched C.S.I.R. on its course of free scientific inquiry, and who had the foresight in the beginning to ensure that official science in Australia would be free from political control, and independent of public service administration.

All of us who have worked in the splendid atmosphere which this created realize the wisdom and vision of the three men who were mainly responsible—a statesman, Viscount Bruce, F.R.S., who as Mr Stanley Bruce was the Prime Minister of Australia 22 years ago when C.S.I.R. was founded; an engineer, Sir George Julius, our first Chairman, who died in office; and a chemist, Sir David Rivett, F.R.S., who left the Melbourne Chair to set the course and take the helm, and who has since tirelessly guided C.S.I.R. on the not particularly easy route of free scientific inquiry.

Our gratitude to these men, and our affection and respect for our leader, are unlimited. We feel that we are the subjects of a grand experiment in official scientific administration and this, perhaps more than anything else, stimulates us to carry the findings of our scientific inquiries a good deal further into the field of practical application than we should be inclined to under any other regime. We are all anxious to prove this experiment a success.

APPENDIX ON HISTORY AND ORGANIZATION OF DIVISION OF BIOCHEMISTRY AND GENERAL NUTRITION

The nucleus of the Division of Biochemistry and General Nutrition was formed when the Council for Scientific and Industrial Research founded the Division of Animal Nutrition in 1927 with the appointment of Professor Thorburn Brailsford Robertson as Chief. Robertson built the central laboratory in the grounds of the University of Adelaide, drew together a small team, and with courage and vision set to work on problems associated with the nutrition of sheep.

Shortly after Robertson died in 1929, Sir Charles Martin, F.R.S., then recently retired from the Directorship of the Lister Institute, London, showed his affection

for Australia by accepting the responsibility of guiding the young Division for a short period. He was Chief of the Division of Animal Nutrition during 1930-32. The wise and inspiring leadership of these men initiated a spirit of research that has held the team together and has directed their efforts. Subsequent to the return of Sir Charles Martin to England in 1932 the Division of Animal Nutrition became part of the newly created Division of Animal Health and Nutrition with Dr L. B. Bull as Chief. In 1944, to simplify administration, the research activities centred at Adelaide were raised to the status of a Division of the Council and the scope was increased by the reference implied in the title—Biochemistry and General Nutrition.

During these changes the course of the research has altered very little—the Division has gradually consolidated into a Research Institute which is concerned broadly with the physiology and biochemistry of ruminants and more particularly with the nutrition of the wool sheep.

The original two-storied Animal Nutrition Laboratory, its annexes which house the two calorimeters, metabolism cages for 20 sheep and subsidiary apparatus, and the four-storied Biochemistry Laboratory now nearing completion which comprise the buildings where the Division is centred, are situated within the grounds of the University of Adelaide. This proximity has fostered a close liaison between the Division and the University; the research staff share alike the amenities and the disciplines of their university colleagues. The Division has a central field station, Glenthorne, of some 600 acres situated about 11 miles from the laboratories, and a series of field stations distributed over the pastoral areas of Australia at sites where nutritional disabilities exist. Experiments conducted on these outlying field stations serve both to clarify the problems and to apply under conditions of station practice the information which results from more intensive studies in the laboratory.

The Division is organized into a number of sections which may be defined loosely as nutritional biochemistry, vital energetics, tissue metabolism, microbiology, analytical and physical chemistry, organic chemistry, animal physiology, plant nutrition, and field stations. The sections flow imperceptibly into one another, the fluidity being fostered to ensure the unhampered intercommunication of ideas essential for successful team work. Long-term individual and collaborative studies which aim at extending knowledge of the nutritional physiology of the sheep constitute the main experimental projects, but on occasions, relatively short-term experimental investigations are undertaken to illuminate the origin and nature of specific nutritional disabilities which climate and terrain impose on grazing ruminants. The guiding principle of all investigations is the simple faith that practical application will flow naturally and freely from a more complete understanding of the underlying phenomena.

The reference of the Division is to conduct research; the research findings are extended to practice in the pastoral industry through the appropriate organizations with which there is a close liaison.

DESCRIPTION OF PLATES 12 TO 18

Plate 12

Typical areas of aeolian soils in Southern Australia where trace element deficiencies prevail

- a The Recent calcareous dunes of the littoral consist essentially of comminuted shell fragments which overlie unconformably and at considerable depth the older soils. The unusually small amount of copper available from these soils limits the development of most plant species. These deficient tracts are overgrown with the calciphile grasses *Bromus madritensis* and *Lagurus ovatus*. Fodder plants which grow there in the absence of manurial dressings of the elements which are short in these soils do not provide enough cobalt and copper to fulfil the physiological requirements of grazing ruminants.
- b The siliceous, solonetz soils of the hinterland support a stunted heath inured by natural selection to this singularly deficient terrain. Growth of other plant species is limited by the deficiencies of phosphorus, zinc and copper which prevail. Ruminants depastured on these tracts develop symptoms of copper deficiency.

Plate 13

The effect of cobalt deficiency on Merino sheep

- a The twin Merino ewes are those referred to in figure 2. The photograph was taken in December 1938 after they had been depastured on cobalt-deficient terrain (see plate 12 a) for 55 weeks during which the ewe on the right, W 6.248, had received the equivalent of 1 mg. Co/day administered *per os* thrice weekly. The other, W 6.241, which was untreated during this period exhibits the symptoms of extreme cobalt deficiency. The oxygen-combining capacity of the blood of the former was at this time 15.5 vol. O₂/100 ml.; that of the latter was 3.5 vol. O₂/100 ml.
- b The morphology of the formed elements in the blood of the sheep above is shown. The film from W 6.241 is typical of the blood dyscrasia which supervenes in the terminal stages of cobalt deficiency; that from W 6.248 is normal.

Plate 14

The effect of copper deficiency on the wool grown by Merino sheep

- a The capacity to impart the crimp which is a feature of Merino wool is lost by the follicles when the animal's copper status is seriously depleted. The three staples illustrated were drawn in three successive years from an area defined on the shoulder of an experimental ewe; the first was grown while the animal was on normal pastures, the other two reflect the deterioration of the animal's copper reserves during two years' grazing on copper-deficient terrain.
- b The ability to impart crimp to the fibre returns abruptly when the copper necessary for this function is provided. The staple illustrated was taken from an experimental ewe which had been dosed with the equivalent of 10 mg. Cu/day while grazing on the deficient pastures on which it had been depleted.
- c When copper-deficient, black-woolled sheep lose their ability to produce melanin and their capacity to impart crimp to their wool fleece. Both functions return immediately when a normal copper status is resumed. The effect of a supplement of copper on the pigmentation and crimp is illustrated. (From experiments of Marston & Lee.)

Plate 15

The effect of different degrees of copper deficiency on the character of Merino wool

When a sheep is confined to pastures which are acutely copper-deficient its reserves of copper are depleted rapidly. The rate of this depletion may be decreased by providing amounts of copper which are insufficient to reinstate a positive copper balance. In this way an ordered series of deficiency states may be imposed. The series of staples shown are from ewes em-

ployed in an experiment (figures 6 and 8) which aimed, by means of a titration, to determine the quantity of copper necessary to reinstate copper equilibrium in sheep depastured on copper-deficient terrain. Each pair of staples was drawn from an identical site in two consecutive years, the first of each were grown while the animal was on normal pastures, the second during the first year of treatment while on deficient pastures. The deficiency lesion appeared in all staples from sheep that received less than the equivalent of 7.5 mg. Cu/day in the supplement. These form a graded series. (From experiments of Marston & Lee.)

Plate 16

Resistance of certain plant species to soil conditions of the calcareous littoral which impose copper deficiency on other species

- a* Experimental cereal crops sown with superphosphate but without copper. Rye corn growing normally under conditions in which oats (var. *mulga*) have developed flagrant lesions of copper deficiency and have failed to head.
- b* Complete failure of certain species of pasture plants (*Phalaris tuberosa*, *Trifolium subterraneum* and *Medicago lupulina*) sown on the calcareous littoral sand manured with ample dressings of superphosphate, potassium sulphate and ammonium sulphate. Left: The sown species have failed to develop and the area is overgrown with the natural species *Lagurus ovatus* and *Bromus madritensis* which have responded to the better nutritional environment provided by the dressings. Right: The natural cover unmanured. (From experiments of D. S. Riceman.)

Plate 17

The establishment of lucerne on the calcareous littoral

- a* Typical shell-sand dunes at Robe, South Australia.
 - Top left:* Area outside rabbit-proof fence denuded of grass-cover by rabbits.
 - Top right:* Typical natural pasture of *Lagurus ovatus* and *Bromus madritensis*.
 - Bottom left:* An experimental crop of lucerne (*Medicago sativa*) established with manurial dressings of copper sulphate.
 - Bottom right:* Experimental quadrates showing effect of the application of copper sulphate.
- b* Quadrate (on bottom right area in *a*) sown May 1938 with lucerne together with manurial dressing of superphosphate and copper sulphate. The lucerne is established. Photographed November 1939.
- c* Quadrate adjacent to *b* sown May 1938 with lucerne together with manurial dressing of superphosphate. In the absence of additional copper the lucerne has failed and the area is dominated by *Lagurus ovatus* and *Bromus madritensis*. Photographed November 1939. (From experiments of D. S. Riceman.)

Plate 18

The development of pastures on the mallee heath soils of the hinterland S.E. South Australia

- a* Area cleared for experimental trials.
- b* and *c*. Experimental pasture of *Trifolium subterraneum*, *Phalaris tuberosa* and some *Medicago sativa* developed on mallee heath soils with manurial dressings of superphosphate, zinc sulphate and copper sulphate, *b* in late spring of second year and *c* in early spring of fifth year after sowing. (From experiments of D. S. Riceman.)

On overpotential and the photovoltaic process at polarized electrodes

By P. J. HILLSON AND E. K. RIDEAL, F.R.S.

The Davy Faraday Laboratory, The Royal Institution

(Received 10 March 1949)

The acceleration of the electrodeposition of hydrogen and oxygen by light of wave-length from 4000 to 2000 Å has been studied on different electrodes. The quantum efficiency of the photoreaction was determined as a function of the wave-length of the incident light and of the steady polarizing current. From these experiments it was concluded that the rate-determining step at these electrodes occurred after the discharge of the ions and that the overpotential corresponded to a stationary concentration of adsorbed atoms or radicals. It was further concluded that an electrode at which hydrogen is being evolved at a sensible rate is covered with a nearly complete layer of adsorbed hydrogen. The potential at which such evolution occurs is determined primarily by the negative potential necessary to displace adsorbed oxygen atoms or water molecules by hydrogen. Calculations of this potential, in the few cases where the necessary heats of adsorption are available, gave values in agreement with those observed experimentally. It was concluded that the rate-determining step on most cathodes, except platinum, was the electrochemical combination of an adsorbed atom with an ion in the double layer, and relations were obtained for the dependence of the overpotential on the current density and electrokinetic potential, which are in fairly good agreement with experiment.

A similar mechanism is tentatively suggested for the electrode reaction at an anode at which oxygen is being evolved.

1. INTRODUCTION

When either hydrogen or oxygen is evolved by electrolysis at metallic electrodes the electrode acquires values in excess of those anticipated for the reversible 'hydrogen' or 'oxygen' electrode respectively. Various alternative mechanisms have been proposed to account for the appearance of these irreversible overpotentials. These have recently been the subject of detailed reviews by Bowden & Agar (1938), Frumkin (1943), Smith (1947), Agar (1947) and Bockris (1948), and it is not necessary to reproduce them all here. We will at this point merely observe that the various hypotheses may be broadly classified into two groups: those in which the rate-determining step in the overall mechanism is the discharge of an ion or the decomposition of a water molecule to give an adsorbed atom or radical, and those in which the rate-determining step occurs after the discharge and the overpotential is due to the resulting stationary concentration of adsorbed atoms or radicals at the electrode surface. The slow-discharge theory (Erdey-Gruz & Volmer 1930; Gurney 1932), in which the rate-determining step is considered to be the transfer of an electron from the cathode to a hydrogen ion or a water molecule falls into the first group, as does the slow proton transfer theory (Eyring, Glasstone & Laidler 1939), in which the rate-determining step is considered to be the transfer of a proton from a water molecule in the double layer to a water molecule adsorbed on the electrode, while the catalytic (Tafel 1905; Hickling & Salt 1942) and electrochemical (Heyrovsky 1925; Okamoto 1937) theories fall into the second.

If the velocity of what is believed to be the rate-controlling process could be altered by some independent method an unambiguous decision as to which is in fact the rate-controlling process could be made. The photovoltaic action of light incident on polarized electrodes has been examined with this object in view, and as a result it has been possible to make such a decision and show that, with certain modifications, the catalytic and electrochemical mechanisms are those which are actually operative at polarized electrodes and to define under what conditions the former or latter process controls the reaction.

Audubert (1929) observed a small photoeffect on both anodically and cathodically polarized electrodes of gold and platinum. The sign of the photo-e.m.f. was such as to cause depolarization of the electrode. Duclaux (1935) observed a decrease in the potential of a tungsten electrode at which oxygen was being evolved when it was illuminated with ultra-violet light. Bowden (1931) observed that both a cathodically polarized mercury electrode and an anodically polarized platinum electrode were slightly depolarized by light of wave-length from 2000 to 4000 Å and made a semi-quantitative determination of the photocurrent (i.e. the extra current required to return the electrode to its original potential) as a function of the electrode material and the wave-length of the incident light. Price (1938) extended the investigation to a variety of electrodes at which hydrogen or oxygen were being evolved and found similar results on all of them. The present work is a more detailed investigation into the effect.

2. EXPERIMENTAL

The electrolysis cell used for measuring the overpotential is shown in figure 1. It was made of 'Pyrex' glass. The quartz window and the electrode being studied were held in place with picein wax. The electrolytic solution, however, never came into contact with this wax or any tap-grease. The cell was partially immersed in a water thermostat at 23° C. The quartz window was warmed with a small lamp, which was switched off during an experiment, to prevent condensation of water vapour on the inside. When a mercury cathode was used the cell was modified, as shown in the inset to figure 1.

The deposition of hydrogen was studied on electrodes of mercury, copper, nickel and silver in a solution of 0.1 N-sulphuric acid, using a $\text{Hg}/\text{Hg}_2\text{SO}_4/0.1\text{ N-H}_2\text{SO}_4$ reference electrode so that no liquid-liquid boundary potentials were established.

The deposition of oxygen was studied on a platinum electrode in 0.1 N-sulphuric acid and on platinum, silver, gold and nickel electrodes in 0.1 N-potassium hydroxide, when a $\text{Hg}/\text{HgO}/0.1\text{ N-KOH}$ reference electrode was used.

The solutions were made from 'Analar' chemicals and twice distilled water. They were electrolyzed for 3 days, using small smooth platinum electrodes, while being stirred with a slow stream of hydrogen (or oxygen, in the case of solutions for use in the experiments on the evolution of oxygen), to remove foreign ions. The solution was placed in the flask shown and a stream of hydrogen (or oxygen), purified from condensable impurities by passing through a liquid-air trap, was passed through the cell and the solution for about 18 hr. The solution was then blown into the cell by means of hydrogen (or oxygen) kept in an aspirator over dilute sulphuric acid.

The electrode, which had been previously cleaned in 40 % nitric acid and washed with distilled water, was prepolarized for 2 or 3 hr. to reduce any surface oxide layer and to bring it to a reproducible condition. Then the electrolyte was renewed.

Under these conditions the hydrogen overpotentials and the oxygen overpotential on gold were reproducible to 2 mV for any one run and to 10 mV for different runs. The other oxygen overpotentials were reproducible to within 20 mV, except on nickel, where variations of the order of 50 mV were observed and on platinum in alkaline solution when a systematic lowering of the overpotential occurred on successive runs.

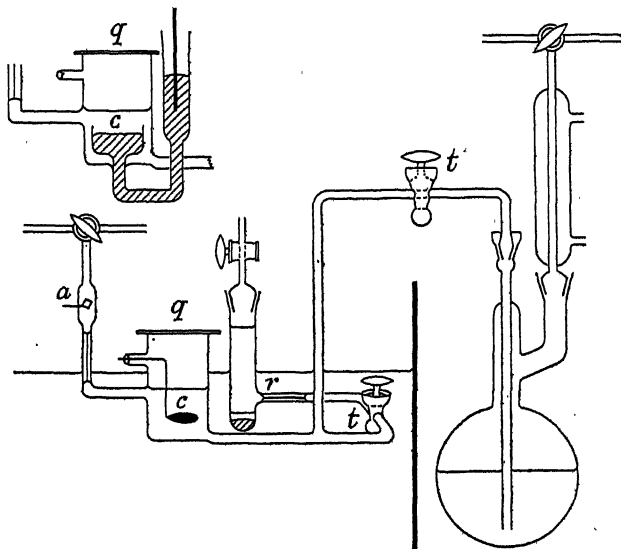


FIGURE 1. The electrolysis cell. *q*, quartz plate; *c*, cathode; *a*, anode; *r*, reference electrode; *t*, greaseless taps.

The electrode was illuminated with monochromatic light, using the large quartz monochromator described by Mitchell & Rideal (1937). The monochromator gave a patch of light about 3 by 7 cm. which was nearly monochromatic and of nearly uniform intensity, from 2 to 10×10^{14} quanta/sq.cm./sec.

The electrode was polarized with a steady current and the overpotential measured. On illumination, the electrode potential fell by 20 to $1000 \mu\text{V}$. The photocurrent could be measured either by increasing the current until the potential returned to its original value or by calculation from the observed change in the overpotential, $\delta\eta$. From the Tafel equation

$$-\eta = a + b \log_{10} i,$$

where $-\eta$ is the overpotential, i the current density and a and b are constants, we obtain

$$\delta i = 2.30i(-\delta\eta)/b.$$

Since the overpotential usually showed a slow drift with time, the latter course was usually adopted, although when both methods could be used good agreement was obtained. The photocurrent was measured at several different current densities for the same degree of illumination, and then the wave-length and intensity of the

incident light were determined separately. The experiments were repeated on the mercury cathode using a 1000 W B.T.H. high-pressure mercury lamp. The light intensities obtained with this lamp were greater by a factor of from 5 to 40, depending on the wave-length, than those obtained previously. The photocurrent, at constant wave-length and polarizing current, were proportional to the light intensity, within the experimental error, showing that the term quantum efficiency, or the number of extra ions discharged per quantum of incident light, is of physical significance when applied to this reaction.

The wave-length of the light was measured with a small 'Beck' ultra-violet spectrometer with a fluorescent screen and the intensity by means of a uranyl oxalate chemical actinometer. 10 ml. of a solution of 0.01 M-uranyl oxalate and 0.05 M-oxalic acid were placed in each of two cells of 4.5 sq. cm. cross-section, approximately the area of the electrodes, and one was illuminated for 2 hr., while the other stood alongside, in the dark, as a blank. 10 ml. of 0.14 N-ceric ammonium sulphate in 10 % sulphuric acid were added each and the excess ceric sulphate determined by titration with 0.005 N-ferrous ammonium sulphate in 10 % sulphuric acid, using phenyl anthranilic acid as indicator. From the amount of oxalate decomposed by the light and the quantum efficiency of the reaction, given by Leighton & Forbes (1930) and Forbes & Heidt (1934), the intensity of the light could be determined. At wave-lengths greater than 3400 Å allowance was made for the incomplete absorption of the light by the oxalate solution, by calculation from the absorption coefficients given by Leighton & Forbes (1930). At wave-lengths less than 3000 Å the oxalate solution was diluted 10 times and was titrated with 0.014 N-ceric ammonium sulphate and 0.001 N-ferrous ammonium sulphate. This had the dual advantage of increasing the titration values in the region where the light intensities were smallest and of preventing the absorption of the light in such a small depth of solution that the quantum efficiency of the reaction would be affected by the inability of the reaction products to diffuse away.

Owing to the difficulty in measuring the potential changes and light intensities, the accuracy of the experiments is probably not better than about 10 % and on the nickel anode is much less.

In order to compare the photoeffect on the different electrodes at the same real current density, the surface areas of the electrodes were measured from the capacities of their double layers (Bowden & Rideal 1928). A 'square-wave' oscillator was built whose frequency was variable from 40 to 250 c./sec. The output was fed on to the control grid of a pentode. The anode current, which could be adjusted to a suitable value by varying the bias of the control and screen grids, passed through the electrolysis cell. The potential between the cathode and the reference electrode was amplified by the amplifier of a cathode-ray oscilloscope and was observed on the screen as a stationary pattern.

In order to reduce 'pick-up' to a minimum, the capacities were determined in a more compact electrolysis cell, using 0.1 N-hydrochloric acid and a Ag/AgCl/0.1 N-HCl reference electrode, which was placed inside the cell close to the cathode. The cell itself was screened by coating it with a conducting layer of 'Aquadag', which was earthed.

Owing to the rather large currents that were employed for the high current pulse (4 mA) there was a noticeable potential drop due to the resistance of the solution. This was apparent as a short vertical line in the potential-time pattern. The capacity of the double layer, and hence the area of the electrode, was inversely proportional to the slope of the potential-time pattern during the charging of the double layer. Since the cell was not adapted to taking a mercury electrode, the capacities were compared with that of a heavily amalgamated copper electrode whose area was assumed to be equal to its apparent area. This assumption was justified by the determination of the ratio real area:apparent area for an amalgamated silver electrode. The value obtained for the ratio was 0.95. The area of the mercury electrode was assumed to be equal to its apparent area.

RESULTS

The values obtained for the ratio real area:apparent area, assuming it to be 1.0 for amalgamated copper, were: amalgamated silver 0.95, gold 1.7, platinum 5.2, nickel 6.9, copper 7.8 and silver 13.3.

The dependence of $\log_{10} \gamma$ on $\log_{10} i$, where γ is the quantum efficiency of the photoreaction and i the polarizing current, in amperes per sq.cm. apparent area, is shown in figures 2 and 3. A series of parallel straight lines is obtained, slope 1.0 to 1.1. Figures 4 and 5 show the dependence of $\log_{10} \gamma$ on $h\nu$, the energy of the incident quanta, measured in eV, at a constant current density of 3.0×10^{-6} amp./sq.cm., corresponding to apparent current densities of $3.0 \mu\text{A}/\text{sq.cm.}$ for mercury, 21 for nickel, 40 for silver and 23 for copper. Straight lines, slope 0.6, are obtained for the silver, nickel

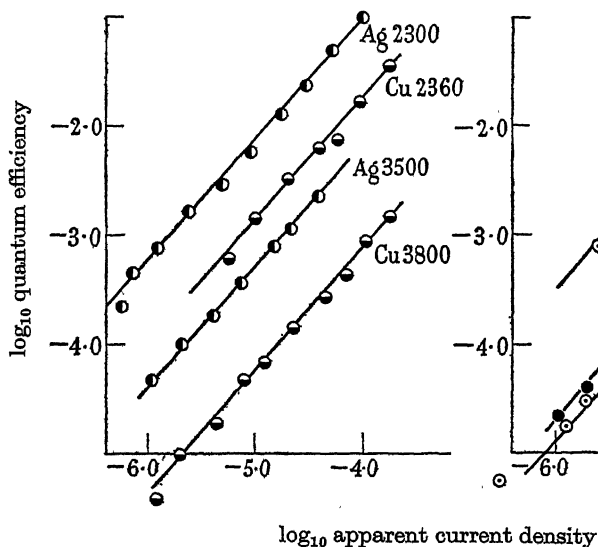


FIGURE 2

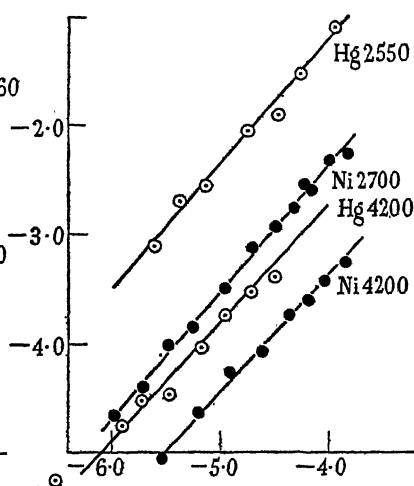


FIGURE 3

FIGURES 2 and 3. The deposition of hydrogen, showing the dependence of $\log_{10} \gamma$ on $\log_{10} i$ at different wave-lengths, given in angstrom units.

and copper electrodes, while the mercury electrode gives a curved line of approximately the same slope. The results may be expressed by the equation

$$\log_{10} \gamma = \text{const.} + \log_{10} i + \beta h\nu / 2.30kT,$$

where $\beta \simeq 1/30$.

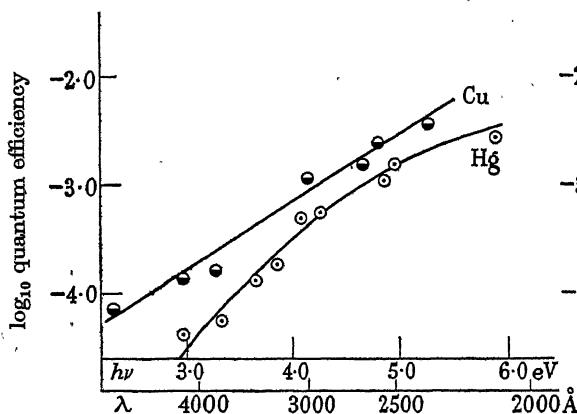


FIGURE 4

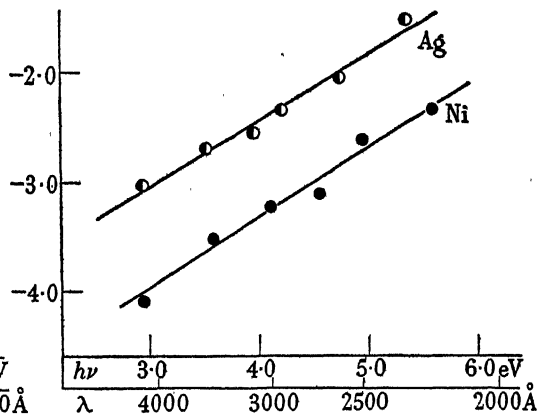


FIGURE 5

FIGURES 4 and 5. The deposition of hydrogen, showing the dependence of $\log_{10} \gamma$ on $h\nu$, the energy of the incident quantum, at a constant polarizing current.

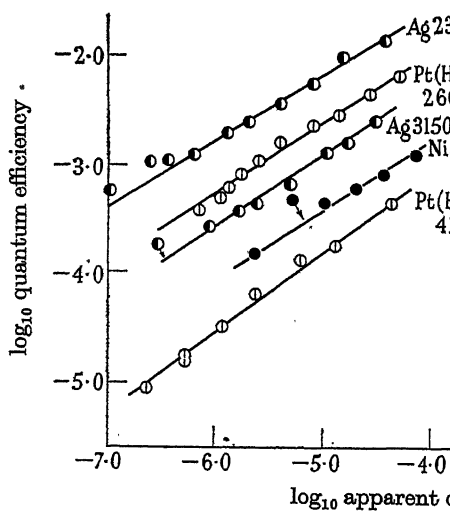


FIGURE 6

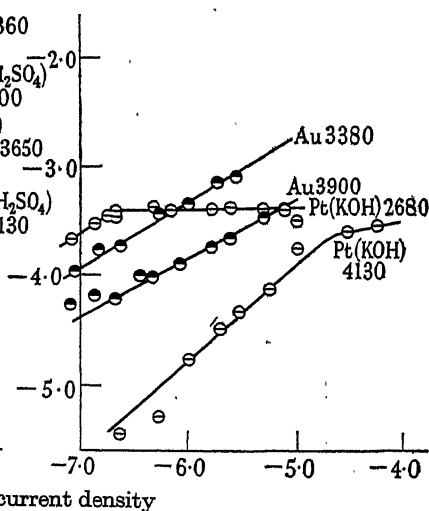


FIGURE 7

FIGURES 6 and 7. The deposition of oxygen, showing the dependence of $\log_{10} \gamma$ on $\log_{10} i$ at different wave-lengths.

There is no threshold frequency in the wave-length range studied, and the constant in the above equation is of the same order on all the metals studied, although the overpotentials, $-\eta$, and the work functions, ϕ , differ considerably. $-\eta$ is 1.29 V with respect to a $\text{Hg}/\text{Hg}_2\text{SO}_4/0.1\text{N-H}_2\text{SO}_4$ electrode on mercury, 1.20 on copper, 1.15 on

silver and 1.10 on nickel. ϕ is 4.52 V for mercury, 4.16 for copper, 4.68 for silver and 5.01 for nickel (Kleine & Lange 1937).

The dependence of $\log_{10} \gamma$ on $\log_{10} i$, at different wave-lengths, is shown in figures 6 and 7, for the deposition of oxygen. A series of parallel straight lines is obtained, slope about 0.6, except on platinum in alkaline solution when γ does not increase beyond a limiting value of 4.4×10^{-4} , as the energy of the quanta and the polarizing current are increased, although the overpotential continues to increase normally with $\log_{10} i$. Figures 8 and 9 show the dependence of $\log_{10} \gamma$ on the energy of the quanta for the deposition of oxygen at a constant current density of 3.0×10^{-7} amp./sq.cm., corresponding to apparent current densities of $2.1 \mu\text{A}/\text{sq.cm.}$ for nickel, 1.55 for platinum, 4.0 for silver and 0.52 for gold. $\log_{10} \gamma$ is again of the same order for all the metals studied and increases as the energy of the quanta is increased. The curves, however, show a change of slope at 4.5 eV on platinum in acid solution and in alkaline solution at 4.3 eV on platinum, 4.5 on silver and 3.5 on gold. The results obtained on nickel were so inaccurate that they only showed that the effect was of the same order of magnitude and the break, if any, could not be located.

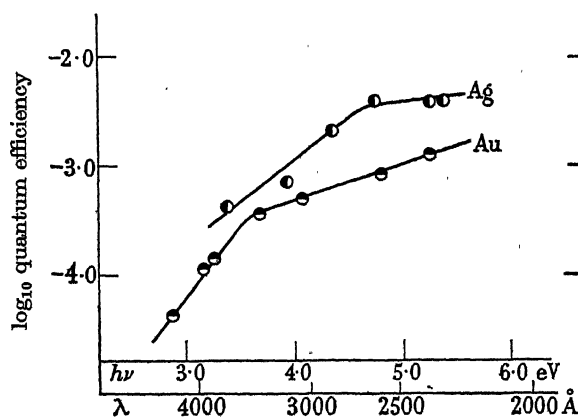


FIGURE 8

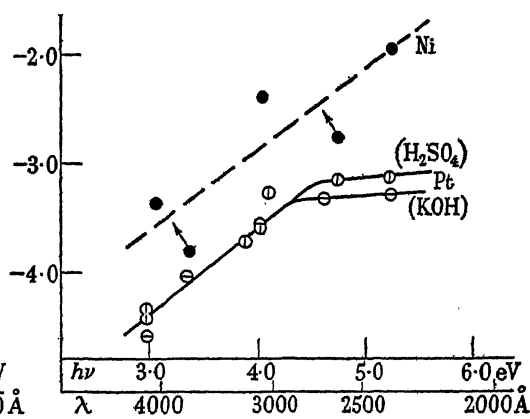


FIGURE 9

FIGURES 8 and 9. The deposition of oxygen, showing the dependence of $\log_{10} \gamma$ on $h\nu$ at a constant polarizing current.

Before discussing the implications of these results on the mechanisms proposed for the electrode reactions, it is necessary to show that they are the result of a true photoeffect and not merely to a heating of the electrode surface by the incident radiation. A simple calculation of the rate of heat input and of the necessary temperature changes involved reveals the fact that, under the most favourable conditions, the surface could not be sufficiently heated above the temperature of the bulk of the electrode and that to heat the whole of it, for an electrode 2 mm. thick, required about 40 sec. Since the potential changes were observed as quickly as the response of the galvanometer allowed, less than a second, the effect due to heating the electrode must be negligible in comparison with the photoeffect. Furthermore, the $\log_{10} \gamma$ - $h\nu$ curves show no sudden breaks over ranges of wave-lengths in which

the reflectivity of the metal, and hence the amount of heat absorbed by the electrode, changes rapidly with the wave-length, e.g. the reflectivity of silver falls from 70 to 10 % between 3500 and 3000 Å.

DISCUSSION

It is evident that if the rate-determining step were the transfer of electrons from the cathode to the ions in the double layer, that photoemission would cause an increase in the current at constant potential. No selective maxima, however, are observed in the $\log_{10} \gamma - h\nu$ curves, as is the case with photocathodes, nor can one interpret the large dependence of the photocurrent on the electrode potential ($\log_{10} \gamma = \text{const.} - 10\eta$) compared with the much smaller dependence on the energy of the incident quanta, measured in eV ($\log_{10} \gamma = \text{const.} + 0.6h\nu$). Further, at an anode, where the rate-determining step is assumed to be the transfer of electrons from the ions in the double layer to the electrode, the photoemission presumably should lead to a decrease in the current and not to the observed reverse.

The slow proton transfer theory has been criticized on other grounds (Butler 1941), but in order to explain the photoeffect it is necessary to assume that the reacting species, water molecules, have a very high absorption coefficient for light when in the adsorbed layer.

We must conclude that these hypotheses in which the discharge rate is the controlling factor are untenable, that the rate-determining step occurs after discharge and that it is accelerated by illumination, whilst the overpotential is due to a stationary concentration of adsorbed atoms or radicals.

We will proceed to examine the catalytic and electrochemical hypotheses in the light of these results.

When clean metallic surfaces are exposed to hydrogen or oxygen, the surfaces are completely, or almost completely, covered with a monolayer of chemisorbed atoms at gas pressures of a fraction of 1 mm. of mercury. We must consequently conclude that when a metal electrode is immersed in aqueous solution it will be completely covered with a chemisorbed monolayer. The constitution of this monolayer must be hydrogen and hydroxyl or oxygen. In this case it is impossible to distinguish experimentally between adsorbed oxygen atoms and hydroxyl radicals and when reference is made, later in the paper, to adsorbed oxygen, it is always possible that it may be hydroxyl.

Since the heat of adsorption of oxygen is usually much greater than that of hydrogen, a monolayer of oxygen atoms is normally formed, the hydrogen being displaced. Langmuir (1916) has shown, for instance, that a tungsten surface rapidly adsorbs water vapour, and its electron emission falls to the value for a tungsten surface on which oxygen has been adsorbed. The heat of adsorption of hydrogen on tungsten is 45 kcal./g.mol. on a bare surface (Beeck 1945), while that of oxygen is about 130 kcal./g.mol. (Roberts 1935*a*), with the result that, as Roberts (1935*b*) showed, gaseous oxygen completely and instantaneously displaces chemisorbed hydrogen as molecular hydrogen. The effect of oxygen poisoning the ortho-para-hydrogen conversion on all metals except platinum shows that this is generally true.

In the case of platinum Oatley (1939) showed that gaseous hydrogen partly displaces chemisorbed oxygen and vice versa, a mixed film being formed in both cases.

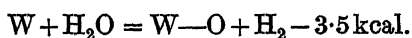
At the negative potentials usually attained during the deposition of hydrogen the chemical potential of the adsorbed hydrogen is much enhanced, the increase corresponding to an increase in pressure by a factor of 3×10^3 for each tenth of a volt in excess of the reversible hydrogen potential, whilst the chemical potential of the oxygen is correspondingly reduced. Under these conditions the chemisorbed oxygen is displaced and replaced by hydrogen. From the high quantum efficiencies observed for the photoreaction at the more negative potentials we conclude that an electrode at which hydrogen is being evolved at a sensible rate is covered with a complete, or nearly complete, layer of chemisorbed hydrogen. The potential at which the hydrogen is evolved, and hence the term a in the Tafel equation,

$$-\eta = a + b \log_{10} i,$$

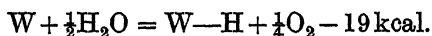
depends primarily on the potential required to displace the chemisorbed oxygen by hydrogen, and hence on the difference in the heats of adsorption of hydrogen and oxygen on the metal.

The experiments of Oatley (1939) indicate that the heats of adsorption of hydrogen and oxygen are comparable, and that even at potentials as low as the reversible hydrogen potential a considerable fraction of a platinum electrode is covered with chemisorbed hydrogen so that hydrogen may be deposited at a sensible rate, and it is possible to use a platinum electrode to measure the reversible hydrogen electrode potential.

Taking Roberts's value of 130 kcal./g.mol. for the heat of adsorption of oxygen on tungsten and 30 kcal./g.mol. from Beeck's figures as the mean heat of adsorption of hydrogen on tungsten, we obtain



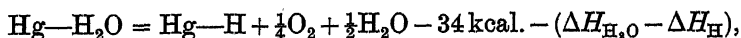
and



Thus the replacement of chemisorbed oxygen by hydrogen, a three-electron transfer, requires an increase of energy of 15.5 kcal./g.atom or 0.67 eV/atom. It should therefore be attained at a potential of -0.22 V with respect to the reversible hydrogen potential. In the case of nickel, where the heat of adsorption of hydrogen is somewhat less (Beeck 1945) while the heat of adsorption of oxygen, as estimated from the heats of formation of the oxides, is probably slightly less too, replacement of oxygen by hydrogen should occur at about the same potential. Both metals, then, should exhibit a low hydrogen overpotential but cannot be used as reversible hydrogen electrodes, as is found.

The correlation observed by Bockris (1947) between the overpotential and thermionic work function may be explained qualitatively on this hypothesis. Emission of electrons from an electrode is equivalent to a surface oxidation. The lower the work function, the easier the oxidation and, in general, the greater the difference in the heats of adsorption of hydrogen and oxygen. Thus, contrary to expectation, on the basis of the slow discharge hypotheses, the lower the thermionic work function the higher the hydrogen overpotential.

Mercury, and possibly other high overpotential metals, may behave differently in that the heats of adsorption of oxygen and hydrogen are so low that water is not dissociated but adsorbed as molecules. Kemball (1946) found that water vapour was adsorbed on to mercury as molecules which were probably associated on the surface. In such a case, the potential at which sensible evolution of hydrogen begins depends primarily on the displacement of water, whose chemical potential does not greatly vary with increasing polarization, by hydrogen:



where $\Delta H_{\text{H}_2\text{O}}$ and ΔH_{H} are the heats of adsorption of water and hydrogen. If these heats are small and of the same order, the required potential is about -1.4 V . Mercury, then, should exhibit a high overpotential for the deposition of hydrogen.

The above calculations are necessarily only approximate. First, because of uncertainties in the measured heats of adsorption, particularly as applied to polarized electrodes, and secondly, because it has been necessary to use the heat of adsorption rather than the free energy and, as Kemball showed for the adsorption of vapours on mercury, these may differ considerably. They do, however, show that the overpotentials to be expected on this view are of the same order as those observed experimentally.

The view that hydrogen is discharged at electrodes which are nearly completely covered with adsorbed hydrogen has previously been objected to on two counts. First, that a change in the electrode potential would involve a change in θ_{H} , the fraction of the surface covered with adsorbed atoms, and the deposition or ionization of these would cause a marked increase in the measured capacity of the electrode, whereas the work of Frumkin (1943) and others, particularly on mercury, has shown that the whole of the capacity of the electrode can be accounted for by the altered charge of the double layer. Secondly, that determinations of the quantity of electricity required to change the electrode potential from that at which hydrogen is being evolved to that at which oxygen is evolved have not detected the presence of hydrogen except on such metals as platinum and palladium. The most recent of such experiments are those of Hickling and others (see Hickling & Taylor 1948), who followed the dependence of potential on the quantity of electricity passed by means of a cathode-ray oscilloscope on electrodes of platinum, gold, nickel and copper. On platinum, they found that reversal of the polarizing current gave a track consisting of three distinct portions. First, a slow change of potential corresponding to the ionization of a monolayer of hydrogen; secondly, a rapid change of potential corresponding to the altered charge of the double layer; and thirdly, a further slow change of potential corresponding to the deposition of a monolayer of oxygen. On the other metals similar results were found for the second and third portions of the track, whereas the first portion was absent.

The conclusion that such experiments prove the non-existence of adsorbed hydrogen on metals other than platinum, makes the tacit assumption that such a layer must be removed by ionization. The work of Roberts (1935*b*) shows that, at any rate *in vacuo*, oxygen displaces hydrogen from tungsten directly as molecular hydrogen, and not as atomic hydrogen or water. We consider that this mechanism is responsible

for the above results and that reversal of the polarizing current first alters the potential of the electrode to that at which oxygen is deposited and that the oxygen atoms then displace the hydrogen without previous ionization, except on platinum where the heat of adsorption of oxygen does not exceed that of hydrogen by an amount great enough to make such a displacement possible. Indeed, unless the surface of the cathodes were nearly covered with adsorbed hydrogen, the deposition of another layer of oxygen prior to the evolution of oxygen would be unnecessary.

The objection on the ground of capacity measurements assumes that a large variation in the potential of the cathode involves a considerable variation in θ_H , the fraction of it covered with hydrogen. If, however, the adsorbed atoms repel one another, by electrostatic forces or otherwise, a very small change in θ_H can involve a large change in the chemical potential of the adsorbed hydrogen (Wang 1938). Trapnell (1949) has recently obtained the following figures, column 2 of table 1, for the equilibrium fraction of a tungsten surface covered with adsorbed hydrogen at different pressures at 0° C.

TABLE 1

p pressure of H_2 (mm. Hg)	fraction of W surface covered by H	$\log i$ equiv. current (amp.)	V equiv. electrode potential (V)
10^{-6}	0.65	4.22	0.242
10^{-4}	0.69	2.17	0.188
10^{-2}	0.73	0.11	0.133

It is seen that even at pressures of a fraction of 1 mm. of mercury, increasing the pressure by a factor of 100 leads to an increase in θ_H of only 4 %. At potentials corresponding to many atmospheres pressure (an overpotential of -0.3 V corresponds to a pressure of 10^{10} atm.), a similar change in chemical potential would involve a change in θ_H which is many times smaller and hence undetectable by the normal capacity measurements.

If a particular electrode potential corresponds to a stationary concentration of adsorbed atoms, the steady cathodic current is equal to the rate of removal of the adsorbed atoms. These can only be removed by diffusion into the metal, by surface combination to give molecular hydrogen (the catalytic mechanism), or by the discharge of and combination of an ion in solution with a surface-adsorbed atom (the electrochemical mechanism). Although hydrogen is known to diffuse into the bulk phase of many metals, it is unlikely that, under steady conditions, the process is rapid enough to constitute a significant fraction of the overall current except possibly on palladium. Diffusion of hydrogen into the bulk of the electrode and of dissolved oxygen from the interior of the electrode to the surface may, however, play a significant part in the slow attainment of a steady overpotential.

In the case of the catalytic mechanism it is clear that, at the degree of saturation of the surface postulated, the rate of combination is dependent on the free energy of the adsorbed atoms rather than on their concentration alone.

We then have the following relationships:

$$-\eta = (G_H - G_O)/F,$$

where G_O and G_H are the free energies of the adsorbed atoms at the reversible and working potentials respectively.

The rate of combination $= i_c/2F$, where i_c is the current due to the catalytic mechanism and

$$\begin{aligned} i_c/2F &= k\theta_H^2 e^{-(E-2G_H)/RT} \\ &= k' e^{2G_H/RT}, \end{aligned}$$

where $k' = k\theta_H^2 e^{-E/RT}$ and E is the energy of activation for surface combination.

$$\begin{aligned} \text{Hence} \quad -\eta &= \text{const.} + (2.30RT/2F) \log_{10} i_c \\ &= \text{const.} + 0.027 \log_{10} i_c \quad \text{at } 0^\circ \text{C.} \end{aligned}$$

This relation is found experimentally for the deposition of hydrogen on platinum, under certain conditions, at low current densities. In the case of the evaporation of hydrogen from tungsten *in vacuo*, where the electrochemical mechanisms cannot be operative, we can take Trapnell's results and place the equilibrium gas pressure equal to an equivalent electrode potential and the rate of evaporation, which is equal to the rate of combination, equal to an equivalent current:

$$V = -(2.30RT/2F) \log_{10} \frac{p}{760} \quad (\text{table 1, column 4}),$$

$$\text{and} \quad \log_{10} i_c = 2.68 + \log_{10} p + \log_{10} (1 - \theta_H) \quad (\text{table 1, column 3}).$$

This gives an expression

$$-V = -0.136 + 0.028 \log_{10} i_c$$

in agreement with that predicted on the catalytic theory.

The relation found on most metals, however, is

$$-\eta = a + (2.30RT/\alpha F) \log_{10} i,$$

where $\alpha = 0.5$ to 0.7 . Hence it must be assumed that on these metals either that the reaction rate is profoundly modified by the presence of a few chemisorbed oxygen atoms or that hydrogen evolution takes place by the other method, the electrochemical mechanism.

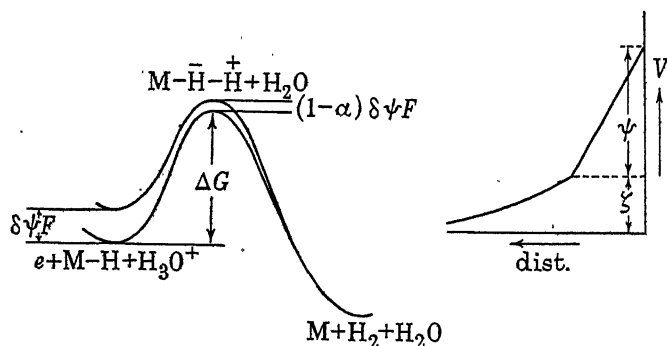
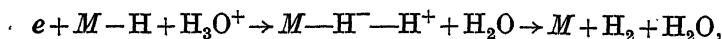
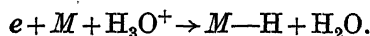


FIGURE 10. The energy and potential barriers at the electrode. The electrode potential V is equal to $\psi + \zeta$, where ψ is the potential across the 'Helmholtz' double layer and ζ is the potential across the 'Gouy' or diffuse portion of the double layer.

In acid aqueous solutions the electrochemical mechanism may be diagrammatically represented as in figure 10 and by the sequence



followed by the rapid reaction



It is evident that for metals which strongly adsorb oxygen, hydrogen evolution is only possible if the potential attains a relatively high cathodic value. Under these conditions the H_3O^+ ions present in the solution will be strongly adsorbed on to the negative $M-H$ surface.

Expressing the rate of reaction, at an electrode potential V , in terms of Eyring's theory of rate processes we obtain

$$i_e/2e = (kT/h) [H] [H_D^+] e^{-\Delta G/RT},$$

where i_e is the current due to the electrochemical mechanism, $[H]$ the concentration of adsorbed atoms and $[H_D^+]$ the concentration of hydrogen ions in the 'Helmholtz' double layer. This is related to $[H_S^+]$, the hydrogen-ion concentration in the bulk of the solution, by the equation

$$[H_D^+] = [H_S^+] e^{-\zeta F/RT},$$

where ζ is the electrokinetic potential or the potential across the diffuse part of the double layer.

Hirota & Horiuti (1938) and Okamoto (1937) have calculated the magnitude of the free energies of the adsorbed atoms and activated complex by statistical methods, but since they did not conceive the whole of the surface to be covered with adsorbed atoms, the values obtained are not valid in this context.

If the potential of the electrode, V , be raised to $V + \delta V$, the free energy of the adsorbed atoms is increased by an amount δVF , while that of the ions in the double layer is decreased by an amount $\delta \zeta F$. Thus the free energy of the reactants is increased by $\delta \psi F$, as shown in figure 10. If the free energy of the adsorbed $M-H^- - H^+$ complex depends on the electric field, ψ , in which it lies and is increased by an amount $(1-\alpha)\delta \psi F$, where $(1-\alpha)$ is a parameter, the rate becomes

$$(i_e + \delta i_e)/2e = (kT/h) [H] [H_S^+] e^{-(\zeta + \delta \zeta)F/RT} e^{-\Delta G/RT} e^{-\alpha \delta \psi F/RT}.$$

Hence $\ln i_e = \text{const.} + \ln [H_S^+] - \zeta F/RT$

and $\ln (i_e + \delta i_e) = \text{const.} + \ln [H_S^+] - (\zeta + \delta \zeta) F/RT - \alpha \psi F/RT.$

Thus $\ln (1 + \delta i_e/i_e) = \delta \ln i_e$
 $= -\delta \zeta F/RT - \alpha \delta \psi F/RT,$

and since $\psi = V - \zeta,$

$$\delta \ln i_e = -\alpha \delta V F/RT - (1-\alpha) \delta \zeta F/RT.$$

Except in the neighbourhood of the electrocapillarity maximum, $V \gg \zeta$, and

$$d(\ln i_e)/dV = -\alpha F/RT$$

or

$$-\eta = a + b \log_{10} i_e,$$

where $b = 2.30RT/\alpha F$, α being a proper fraction, and the value of a , as explained above, depends primarily on the difference in the heats of adsorption of oxygen and hydrogen.

In dilute acid solution, in the absence of foreign electrolytes,

$$\zeta = \text{const.} + (RT/F) \ln [H_3^{+}].$$

If ΔG is constant at a constant overpotential, then on high overpotential metals, where the metal is far from the electrocapillarity maximum, $\log_{10} i_e$ is independent of $[H^+]$. Levina & Sarinsky (1937) found this to be true for the deposition of hydrogen on mercury in dilute hydrochloric acid solution.

At constant current density and hydrogen-ion concentration

$$\delta V + (1 - \alpha) \delta \zeta = 0.$$

On mercury, where α is found experimentally to be equal to 0.5, then

$$V + \zeta = \text{const.},$$

an expression first derived by Frumkin (1933) from other considerations.

Illumination leads to the activation of some of the adsorbed hydrogen atoms. Winch (1931), on the basis of some experiments on the 'outgassing' of photocathodes by exposure to ultra-violet light, concluded that the actual activation may be done by the ejected photo-electrons. In any case, the number of atoms activated per second is proportional to $I[H]$, where I is the intensity of illumination. The activated atoms quickly reattain thermal equilibrium with the substrate, but a fraction are able to react before they are deactivated.

If the average fraction of the energy acquired that remains when the atoms are able to react is β , the

$$\begin{aligned} \text{rate of the photoreaction} &= i_p/2e \\ &= (kT/h) [H^+] [H_D^+] e^{-\Delta G/RT} e^{\beta h\nu/kT}, \end{aligned}$$

where $[H^+]$ is the number of atoms activated by the light per second and is equal to $kI[H]$.

So that

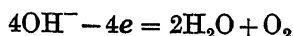
$$i_p = i_e k I e^{\beta h\nu/kT},$$

and

$$\log_{10} \gamma = \log_{10} i_p/I = \text{const.} + \log_{10} i_e + \beta h\nu/2.30kT,$$

where the constant is of the same order on all metals. This is the expression found experimentally for the dependence of $\log_{10} \gamma$ on $\log_{10} i$ and $h\nu$.

At an anode at which oxygen is being evolved there can be no question of the surface being covered with a mixture of chemisorbed hydrogen and oxygen, since the only metals on which it may be assumed that the heats of adsorption of hydrogen and oxygen are comparable, platinum and palladium, show high oxygen overpotentials. Consequently it is assumed that oxygen is discharged on a metal oxide layer that is at least one molecule thick and on some metals, e.g. copper, may attain a considerable depth after prolonged polarization. While the mechanism of the reaction



is probably more complex than the reaction at the cathode, it is assumed that the overpotential is due to a stationary concentration of hydroxyl radicals (or oxygen atoms) adsorbed on the oxide layer. The rate of deposition of these radicals is then equal to their rate of removal by a catalytic or electrochemical mechanism.

Owing to the uncertainty as to the nature of the adsorbed particles and their concentration it is not possible to formulate an exact expression for the rate of deposition of oxygen, but it is probably significant that the transition elements, iron, cobalt and nickel, which have a high affinity for oxygen, have a low oxygen overpotential, while the noble metals, platinum, and gold, have high overpotentials.

If G_{OH} and G_{O} are the free energies of the adsorbed hydroxyl radicals at the working and reversible potentials of the electrode respectively

$$\eta F = G_{\text{OH}} - G_{\text{O}} = f(\theta_{\text{OH}}),$$

and if the radicals are mainly removed by catalytic combination,

$$\begin{aligned} \text{rate of combination} &= i_c/2F \\ &= k\theta_{\text{OH}}^2 e^{-(E-2f(\theta_{\text{OH}}))/RT} \end{aligned}$$

and

$$\eta = a' + b' \log_{10} i_c,$$

where b' is not necessarily constant at different polarizations.

The primary action of the light is supposed to be the decomposition of a molecule of the surface oxide to give an activated oxygen atom which may react with one of the adsorbed radicals before it is deactivated and returns to its place in the oxide lattice.

If $Nh\nu > E'$, the energy required for the reaction



to give a free oxygen atom, the stationary concentration of free oxygen atoms produced by the illumination, $[O^+]$, is proportional to the light intensity, I , and if β' is the average fraction of the excess energy, $h\nu - E'/N$, which they retain until they are able to react

$$\begin{aligned} i_p/2F &= k\theta_{\text{OH}}[O^+] e^{-(E-f(\theta_{\text{OH}})-\beta'(Nh\nu-E'))/RT} \\ &= k'I\sqrt{i_c} e^{\beta'h\nu/kT} \end{aligned}$$

and $\log_{10} \gamma = \log_{10} i_p/I = \text{const.} + \frac{1}{2} \log_{10} i_c + \beta'h\nu/kT \cdot 2.30.$

If, on the other hand, $Nh < E'$, the stationary concentration of free oxygen atoms, $[O^+]$, is governed by the rate at which oxygen atoms activated inside the lattice can diffuse to the point of reaction and may be expressed as

$$[O^+] = k'I e^{-\rho(E'-Nh\nu)/RT},$$

where ρ is an undefined parameter.

Then $\log_{10} \gamma = \text{const.} + \frac{1}{2} \log_{10} i_c + (\beta' + \rho)h\nu/2.30kT.$

Similar relations are obtained if it is assumed that the radicals are removed primarily by the electrochemical mechanism, provided that it is assumed that the primary action of the light is to decompose a molecule in the oxide layer.

At a constant wave-length, the slope of the $\log_{10} \gamma$ - $\log_{10} i$ curves would be expected to be 0.5, while the value found experimentally is about 0.6. The limiting value of 4×10^{-4} found for γ on platinum in alkaline solution if attributed to the limiting quantum efficiency for the primary photoreaction, $\text{PtO} + h\nu = \text{Pt} + \text{O}$, presumably due to the low absorption coefficient of an oxide layer that was probably only one molecule thick.

At a constant current density, the $\log_{10} \gamma$ - $h\nu$ curves would be expected to show a fairly marked change of slope when $Nh\nu = E'$. E' is calculated from thermal data to be: PtO 99 kcal./g.mol. (4.3 eV/molecule), Au_2O_3 77 kcal. (3.4 eV), Ag_2O 89 kcal. (3.9 eV) and NiO 139 kcal. (6.1 eV). Breaks were observed in the experimental curves at 4.5 and 4.3 eV on platinum in acid and alkaline solution respectively and at 3.5 eV on gold, in very good agreement with the theoretical values, while the value of 4.5 eV obtained on silver, under less favourable experimental conditions, is in fair agreement with the theory.

REFERENCES

- Agar, J. N. 1947 *Ann. Rep. Chem. Soc.* **44**, 7.
 Audubert, R. 1929 *C.R. Acad. Sci., Paris*, **189**, 800.
 Beeck, O. 1945 *Rev. Mod. Phys.* **17**, 61.
 Bockris, J. O'M. 1947 *Nature*, **159**, 539.
 Bockris, J. O'M. 1948 *Chem. Rev.* **43**, 525.
 Bowden, F. P. 1931 *Trans. Faraday Soc.* **27**, 505.
 Bowden, F. P. & Agar, J. N. 1938 *Ann. Rep. Chem. Soc.* **35**, 90.
 Bowden, F. P. & Rideal, E. K. 1928 *Proc. Roy. Soc. A*, **120**, 59.
 Butler, J. A. V. 1941 *J. Chem. Phys.* **9**, 279.
 Duclaux, J. P. E. 1935 *C.R. Acad. Sci., Paris*, **200**, 1838.
 Erdey-Gruz, T. & Volmer, M. 1930 *Z. phys. Chem. A*, **150**, 203.
 Eyring, H., Glasstone, S. & Laidler 1939 *J. Chem. Phys.* **7**, 1053.
 Forbes, G. S. & Heidt, L. J. 1934 *J. Amer. Chem. Soc.* **56**, 2363.
 Frumkin, A. 1933 *Z. phys. Chem. A*, **164**, 121.
 Frumkin, A. 1943 *Acta Physiochim. U.R.S.S.* **18**, 23.
 Gurney, R. W. 1932 *Proc. Roy. Soc. A*, **134**, 137.
 Heyrovsky, J. 1925 *Rec. Trav. chim. Pays-Bas*, **44**, 499.
 Hickling, A. & Salt, P. W. 1942 *Trans. Faraday Soc.* **38**, 474.
 Hickling, A. & Taylor, D. 1948 *Trans. Faraday Soc.* **44**, 262.
 Hirota, K. & Horiuti, J. 1938 *Bull. Chem. Soc. Japan*, **13**, 228.
 Kernball, C. 1946 Dissertation, Cambridge University.
 Kleine, O. & Lange-Erlangen, E. 1937 *Z. Elektrochem* **43**, 570.
 Langmuir, I. 1916 *J. Amer. Chem. Soc.* **38**, 2221.
 Leighton, W. G. & Forbes, G. S. 1930 *J. Amer. Chem. Soc.* **52**, 3139.
 Levina, S. & Sarinsky, W. 1937 *Acta Physiochim. U.R.S.S.* **6**, 491.
 Mitchell, J. S. & Rideal, E. K. 1937 *Proc. Roy. Soc. A*, **159**, 206.
 Oatley, C. W. 1939 *Proc. R. Phys. Soc. Lond.* **51**, 318.
 Okamoto, G. 1937 *J. Fac. Sci. Hokkaido Univ.* **iii**, 2, 115.
 Price, G. 1938 Dissertation, Cambridge University.
 Roberts, J. K. 1935a *Proc. Roy. Soc. A*, **152**, 464.
 Roberts, J. K. 1935b *Proc. Roy. Soc. A*, **152**, 477.
 Tafel, J. 1905 *Z. Phys. Chem.* **50**, 641.
 Smith, J. W. 1947 *Sci. Progr.* **35**, 675.
 Trapnell, B. M. W. Unpublished results.
 Wang, J. S. 1938 *Proc. Camb. Phil. Soc.* **34**, 224.
 Winch, R. P. 1931 *Phys. Rev.* **38**, 321.

Observations on the turbulent fluctuations of a tidal current

BY K. F. BOWDEN AND J. PROUDMAN, F.R.S.
Department of Oceanography, University of Liverpool

(Received 8 April 1949)

Records have been obtained of fluctuations in the speed of the tidal current in the Mersey estuary, using a current meter in a stand on the bottom, and compared with other records taken with the meter suspended freely at various depths. The fluctuations covered a wide range of periods but could be separated into two main types: 'short period', having periods of the order of a few seconds, and 'long period', with periods from 30 sec. to several minutes. The amplitudes, periods and auto-correlation of the short-period fluctuations have been examined in some detail, and it is concluded that the fluctuations observed near the bottom are evidence of the turbulence associated with bottom friction. It is believed to be the first time that the presence of turbulent velocity fluctuations of this time-scale in the sea has been established experimentally. The long-period fluctuations resemble those found in previous investigations and show features consistent with their being turbulent in origin also, although turbulence of the time-scale involved in their case would probably be mainly horizontal.

1. ORDER OF MAGNITUDE

Let $[\]$ denote a mean value taken over a few minutes at a fixed point. Suppose that the direction of the mean current is in a fixed horizontal line, and let U denote the speed of this mean current. Let $U+u$, v , w denote components of the total current at any instant, so that w is vertical. Then

$$[u] = [v] = [w] = 0,$$

and the Osborne-Reynolds expression for the force of internal friction is

$$\rho[uw]$$

per area, ρ denoting the density of the water.

This paper is concerned with measurements of $U+u$ and a preliminary estimate of the magnitudes of $[u^2]$ to be expected may be obtained from the empirical formula

$$0.0025\rho U^2 \tag{1.1}$$

per area, for the force of friction on the bottom. Near the bottom we have two expressions for the force of internal friction, and equating them gives

$$[uw] = 0.0025U^2. \tag{1.2}$$

On writing $[w^2] = \lambda^2[u^2]$, $\frac{[uw]}{\{[u^2][w^2]\}^{\frac{1}{2}}} = r$,

so that λ indicates the ratio of vertical to longitudinal turbulence, and r is the correlation coefficient between these two components of turbulence, we have

$$[u^2] = \frac{0.0025}{\lambda r} U^2. \tag{1.3}$$

Now, for a given value of λ , the ratio

$$[u^2]^{\frac{1}{2}}/U$$

will increase as the value of r decreases. If we suppose that

$$\lambda = 1, \quad r = 0.3,$$

which is in accordance with the figures given by Kármán (1935), based on small-scale experiments on flow between parallel plates, we have

$$[u^2]^{\frac{1}{2}}/U = 0.09. \quad (1.4)$$

When $U = 100$ cm./sec., or approximately 2 knots, the equation (1.4) gives

$$[u^2]^{\frac{1}{2}} = 9 \text{ cm./sec.}$$

2. PREVIOUS INVESTIGATIONS

Thorade (1931) published a diagram showing a record taken in the River Elbe in June 1926, by means of the current meter designed by H. Rauschelbach. This meter recorded both speed and direction of current about every 10 sec. It was held about 2 m. below the bottom of a pontoon, which itself was about 1 m. below the surface of the river. The depth of water under the meter was about 5 m. The diagram extended over about 17 min. and showed an oscillation with a period of about 5 min. and a speed amplitude of about 10 cm./sec.

In 1934 Thorade published an analysis of some observations taken from the *Poseidon* in the international Kattegat Expedition of August 1931. Three Rauschelbach current meters were worked simultaneously from the same ship, and the observations extended over about a week. Within an interval of 3 min. the speed of the current varied over 10 cm./sec., about a mean value of 25 cm./sec. Taking 1 min. mean values of the components, well-marked oscillations with periods of 3 to 5 min. and with amplitudes of 4 to 6 cm./sec. were revealed. Thorade then considered the difference between the simultaneous readings of two instruments. This eliminated the effect of the translatory motion of the ship and allowed the effect of the changing direction of the ship to be calculated. The remaining part of the difference was split up into a quick variation and a slow variation, the latter being given by 3 min. mean values. Thorade found that the quick variation increased in amplitude with the magnitude of the surface waves, but was independent of the mean speed of the current; he concluded that it was due to the rolling of the ship. He found the slow variation to be independent of the surface waves, but to depend partly on the speed of the current and partly on the rate of mixing of the water as revealed by observations on salinity.

In 1936 Doodson (1940) designed and built a current meter which gives a photographic record of all fluctuations of speed whose period is greater than 1 sec., and this meter has since been used by the staffs of the Liverpool University Oceanography Department and the Liverpool Observatory and Tidal Institute. During two weeks in 1938 the meter was hung freely from a boat, the *Zephyr*, anchored about 4 sea miles from the south coast of the Isle of Man, where the current reaches about 150 cm./sec. at spring tides. For a number of records the meter was below the direct influence of the surface waves. An inspection of the records showed at once the presence of two oscillations, one with a period of about 3 sec. and the other with

periods of the order of 1 min. The quicker oscillation occurred all the time, while the slower oscillation occurred only in groups. The *Zephyr* rolled for much of the time, and its period of rolling was about 3 sec. Though experiments made in harbour showed that artificial rolling produced no fluctuations on the records, yet it was decided not to publish the records or an analysis of them.

In 1946 the investigation was continued by attaching the current meter to one of the anti-aircraft forts in Liverpool Bay. Records were taken during two weeks, but on each occasion there was so much surface wave that the investigation became a study of surface waves (Bowden 1947).

Mosby (1946) used a current meter, consisting of twelve sets of revolving cups, to measure the current at twelve different levels near the bottom of Alvaerstrømmen, near Bergen. He discovered oscillations with periods of about 5 min.

3. OBSERVATIONS IN THE MERSEY

The observations studied in the present paper were made in August 1948 from an anchored boat, the *William Herdman*, in the River Mersey. The *William Herdman*, a 62 ft. motor fishing vessel, is the property of the University of Liverpool and is managed by R. J. Daniel. Doodson's meter was again used, and all the members of the Oceanography Department took part in the work.

The *William Herdman* occupied two positions in the River Mersey:

(1) 3 to 6 August 1948, she was moored to a buoy off Bromborough in a depth of 10 m. at L.W.S. (low-water spring tides). The width of the river there is 1.54 km. at L.W.S. and 2.06 km. at H.W.S., and the maximum depth in the section is 10 m. at L.W.S.

(2) From 16 to 20 August 1948, she was at anchor in a position 0.5 km. below Rock Ferry pier, i.e. about 2 km. down river from the former position in a depth of 17 m. The river is 1.0 km. wide at L.W.S. and 1.82 km. at H.W.S. in this position and the maximum depth is 18 m. at L.W.S.

The range of the tide in the Mersey is about 9 m. at spring tides. Both the positions were to the west side of the centre of the shipping channel. The nature of the bed is mainly stone and shell in this part of the river, with mud-banks near the sides.

In normal use the Doodson current meter is suspended through the swivel, by its electric cable. For the measurements near the bottom, a stand was constructed of $1\frac{1}{4}$ in. diameter iron tubing, joined at the corners by brass castings (see figure 1). It consisted of four vertical tubes, each 8 ft. high, spaced 5 ft. apart by two horizontal frames, one at the top and the other from 1 to 3 ft. above the pointed feet at the lower ends of the vertical tubes. The lower frame was extended for $2\frac{1}{2}$ ft. to the front and to the rear, with vertical fins at the rear end to orientate the stand in the current when being lowered, and horizontal boards at both ends to obviate any tendency for the stand to tilt over on the bottom due to the pressure of the current. The whole of the lower frame could be adjusted in height, according to the nature of the bottom, so that the vertical feet could sink in as far as the frame. From the top corners, four tubes converged upwards to a casting carrying an eye-bolt to which the supporting wire was shackled. The meter was suspended inside the stand by a wire from the top,

the length of which could be altered to adjust the position of the meter vertically. This stand was built in the workshop of the Liverpool Observatory and Tidal Institute, under the supervision of A. T. Doodson.

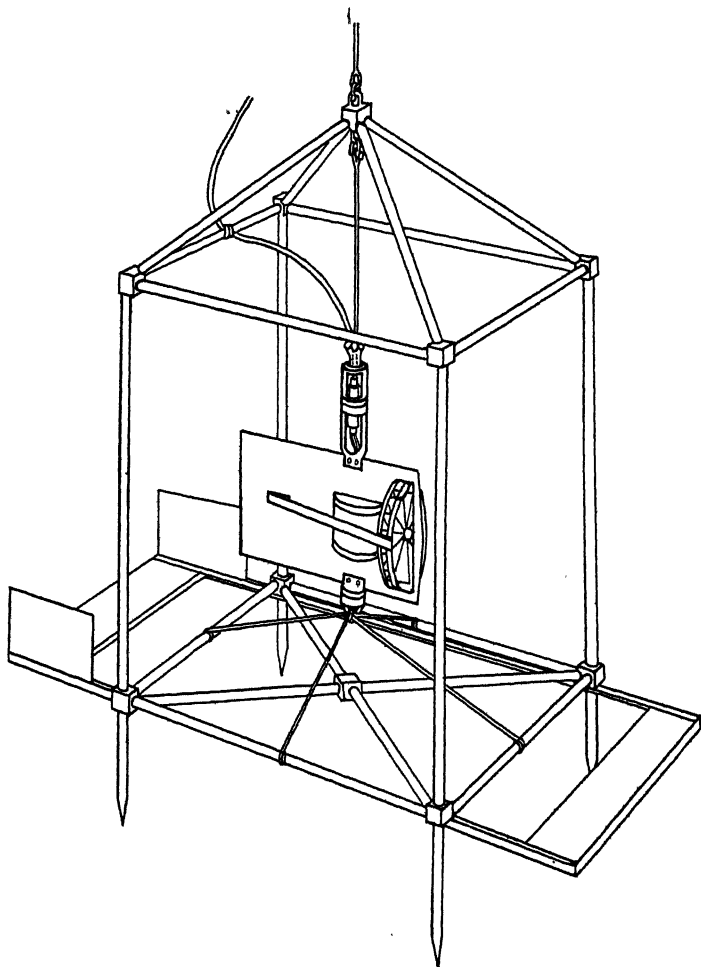


FIGURE 1. Stand for supporting the Doodson current meter on the bottom.

The stand was lowered over the side of the boat by a steel wire leading from a power-operated winch over a snatch-block on the derrick. The electric cable was kept slack and free of the wire and led to a hand-operated winch in the hold.

The stand was usually lowered to the bottom soon after slack water and raised before the next slack water, the wire being kept slack while the stand was on the bottom. As the boat was moored or anchored from the bow only, it was necessary to wait after slack water until the boat was streaming in the current and in a fairly steady position. Similarly, it was necessary to raise the stand before the boat started to swing at the following slack water. When the effects of wind and tide on the boat were in opposing directions, the movements of the boat were often considerable, and it was difficult to prevent the wire becoming taut and causing the stand to tilt.

A device was fitted in the meter giving an indication on board when it was tilted through an angle of more than about 20° , and records during which the stand was believed to have been tilted were rejected.

The recording gear, housed in the hold, gave a record of the current on photographic paper 60 mm. wide, moving at a speed of about 4 cm./min. Time marks were put on the record automatically every 30 sec.

In taking measurements at depths not near the bottom the meter was suspended from the wire directly, without the stand, the other arrangements being as before.

Omitting records which were invalidated by tilting or other faults, fifty-five successful current records were obtained, each of about 30 min. duration. Of these, thirty-seven were taken near the bottom with the meter in the stand and the remaining eighteen at various depths with the meter suspended freely.

The Favé gauge, modified for recording wave pressures, as previously described (Bowden 1947), was laid on the bottom and records obtained simultaneously with the current records on several occasions. Since the recording disk made one revolution in 10 min., provision was made for switching the driving motor in the gauge on and off from the boat, through an electric cable. The gauge was laid shortly after slack water and left down until near the following slack water. In that time three 10 min. traces at intervals of about 2 hr. could be obtained on one disk. Owing to the change in depth of water due to tide, the three traces were separated radially and could easily be distinguished.

4. MEASUREMENT OF RECORDS

The fluctuations on most of the current records were very variable, but it seemed possible to divide them into two main types: (a) fluctuations with periods of the order of a few seconds, which were superimposed on (b) fluctuations with periods of from 30 sec. to several minutes. The term 'period' is used here to indicate the time scale of the fluctuations and does not imply the existence of a series of regular oscillations.

The mean amplitude of the short-period fluctuations was estimated for each minute interval of the record, and the mean period estimated from the number of peaks in the same interval. To measure the long-period fluctuations, a tracing was made of each record, smoothing out the short periods. The mean amplitude and period of the long-period fluctuations were then estimated in a similar way, but using 10 min. intervals. Elimination of the long-period fluctuations gave the mean current. Unless there was a marked change in the character of the fluctuations during a 30 min. record, mean values were deduced for the whole record.

The measured mean values for the records taken with the current meter in the stand on the bottom are given in tables 1 and 2, referring to the position at Bromborough Buoy and off Rock Ferry respectively. The corresponding results with the meter suspended freely at various depths are given in table 3. Denoting the mean current by U , the mean amplitude of the short-period fluctuations by u_1 , and that of the long-period fluctuations by u_2 , the ratios u_1/U and u_2/U are also given in the tables. Figure 2 shows examples of current records (a) near the surface and (b) near

TABLE 1. CURRENT RECORDS AT BROMBOROUGH BUOY WITH METER IN STAND ON BOTTOM

1	2	3	4	5	6	7	8	9	10	11	12	13	14			
record no.	date August 1948	time (central) B.S.T. h. m.	tide	depth of water (m.)	height above bottom (m.)	mean current U (cm./sec.)	short-period fluctuations		u_1/U	long-period fluctuations		remarks				
							T_1 (sec.)	period amplitude u_1 (cm./sec.)		T_2 (sec.)	period amplitude u_2 (cm./sec.)					
16	8	13 55	ebb	17.0	1.4	65	5.0	5.8	0.089	59	5.9	0.092	—			
17	5	14 44		16.1		57	4.4	6.7	0.118	54	6.5	0.114		—		
18	5	15 23		14.9		47	4.9	6.8	0.145	60	5.3	0.113			—	
19	5	16 07		13.7		44	4.9	6.7	0.152	91	5.3	0.121				—
20	5	16 53		12.5		30	5.3	4.7	0.157	130	3.7	0.123				
21a	5	20 45	flood	11.9	57	3.5	2.7	0.047	54	1.9	0.033	6 min. only. P				
21b	5	21 29		13.4	81	5.1	5.5	0.068	210	6.7	0.083		10 min. only. P			
22	6	02 26	ebb	17.0	67	4.8	4.5	0.067	42	6.2	0.093	—				
23	6	03 55		14.6	48	4.4	6.3	0.132	55	6.4	0.134		—			
24	6	04 36		13.4	35	5.0	4.5	0.128	59	3.9	0.112			—		
25	6	05 14		12.2	25	5.2	2.4	0.096	96	1.7	0.068				—	

Note. The letter P in the 'remarks' column indicates that the period P is in minutes.

Note. The letter P in the 'remarks' column indicates that a pressure-gauge record was obtained at the same time as the current record.

TABLE 2. CURRENT RECORDS AT ANCHOR OFF ROCK FERRY WITH METER IN STAND ON BOTTOM

1	2	3	4	5	6	7	8	9	10	11	12	13	14
record no.	date August 1948	time (central) B.S.T. h. m.	tide	depth of water (m.)	height above bottom (m.)	mean current U (cm./sec.)	short-period fluctuations		u_1/U	long-period fluctuations		remarks	
							T_1 (sec.)	u_1 (cm./sec.)		T_2 (sec.)	u_2 (cm./sec.)		
28	16	12 47	ebb	22.9	1.1	29	6.0	3.8	0.131	122	2.9	0.100	—
29	16	13 26		22.2		23	5.6	2.3	0.100	(80)	0.9	0.039	—
30	16	14 15		21.3		25	(6.6)	2.0	1.080	(210)	1.1	0.044	—
31	16	15 06		20.4		21	—	—	—	—	—	—	occasional short-period pulses
32	16	18 49	flood	21.0	1.1	38	4.2	4.6	0.121	280	3.5	0.092	—
33	16	19 28		21.9		56	5.2	5.1	0.091	124	5.5	0.098	P
34	16	20 08		23.1		55	4.4	5.4	0.098	77	4.3	0.078	—
35	16	20 57		24.1		56	(2.0)	10.9	0.195	57	7.1	0.127	violent fluctuations of usually short period, P
36	16	21 37	24.6	43	2.3	13.1	0.305	46	6.7	0.156			
37	16	23 55	ebb	24.3	1.1	23	4.0	2.4	0.104	72	2.4	0.104	—
38	17	00 39		23.8		39	—	5.4	0.138	—	4.6	0.118	8 min. only
39	17	01 36		22.6		29	(5.2)	3.2	0.110	150	1.0	0.035	12 min. only
40	17	02 49		21.0		28	(5.7)	2.5	0.089	—	—	—	long-period fluctuations negligible
41	17	03 32	20.1	23	(5.0)	1.5	0.065	—	—	—			
42	17	08 02	flood	22.2	1.1	55	5.0	7.6	0.138	61	7.9	0.144	—
43	17	08 45		23.2		60	3.9	7.2	0.120	47	8.7	0.145	—
44	17	09 35		24.4		60	4.3	6.9	0.115	57	6.5	0.108	—
45	17	10 17		25.0		51	4.8	5.1	0.100	74	6.0	0.118	—
46	17	10 59		25.0		27	3.8	4.9	0.181	51	5.3	0.196	20 min. only
47	17	13 18	ebb	24.7	0.6	39	6.3	7.3	0.187	71	5.4	0.138	P
48	17	14 15		23.1		27	6.3	5.3	0.196	103	3.1	0.115	—
49	17	15 05		22.3		21	7.2	3.7	0.176	94	2.5	0.119	P
50	17	15 46		21.3		13	(6.0)	1.9	0.146	—	0.7	0.054	fluctuations negligible, P
51	17	16 40	20.4	(7)	—	—	—	—	—	—			
52	17	19 38	flood	21.9		45	4.5	3.9	0.087	87	3.3	0.073	—
54	18	16 09	ebb	18.9		34	5.3	2.5	0.073	(210)	1.8	0.053	—

TABLE 3. CURRENT RECORDS WITH METER SUSPENDED FREELY

1	2	3	4	5	6	7	8	9	10	11	12	13	14
record no.	date August 1948	time (central) B.S.T. h. m.	tide	depth of water (m.)	distance below surface (m.)	mean current U (cm./sec.)	period T_1 (sec.)	amplitude u_1 (cm./sec.)	short-period fluctuations u_1/U	period T_2 (sec.)	amplitude u_2 (cm./sec.)	long-period fluctuations u_2/U	remarks
27	6	18 26	ebb	—	1.8	38	4.5	1.1	0.029	(156)	1.1	0.029	sea 0-1
55	19	08 35	flood	19.2	3.4	65	2.9	1.3	0.020	228	2.3	0.035	sea 0
56	19	09 17		20.4	6.4	83	3.2	5.3	0.064	64	4.7	0.056	
57	19	10 05		22.0	9.5	88	3.7	7.0	0.080	71	4.1	0.046	
58	19	10 47		22.9	12.8	77	3.7	6.9	0.090	67	4.7	0.061	
59	19	11 24		23.5	16.5	64	4.2	6.3	0.098	53	5.3	0.083	—
61	19	14 09	ebb	22.9	3.0	101	3.5	7.9	0.078	91	2.7	0.027	sea 2. P
62	19	14 48		22.0	6.1	117	3.6	7.7	0.066	—	(<1.2)	—	
64	19	16 09		20.1	11.9	70	4.5	5.6	0.080	—	4.0	0.057	
66	19	17 29		18.3	14.6	40	5.2	5.9	0.147	86	6.5	0.162	
69	19	23 15	flood	23.5	3.7	92	(2.2)	13.5	0.147	52	8.1	0.088	sea 1. P P 17 min. only
68	19	22 25		22.6	5.5	101	3.3	9.7	0.096	40	4.1	0.041	
70	20	00 06		24.4	11.0	85	3.7	7.8	0.092	126	7.8	0.092	
67	19	20 38		19.2	11.6	62	3.1	4.2	0.068	165	5.5	0.089	
76	20	06 45	ebb	17.4	7.3	27	4.8	1.9	0.070	156	3.4	0.126	5 min. only 12 min. only
73	20	03 25		21.6	11.3	101	4.4	4.7	0.047	—	—	—	
74	20	04 18		20.1	14.9	107	5.5	5.5	0.051	60	4.1	0.038	
71	20	01 45		23.5	19.5	87	4.2	4.5	0.052	126	5.4	0.062	

the bottom. The very short period, about 1 sec., is instrumental, corresponding to the natural period of the recording galvanometer.

Records of the pressure on the bottom were obtained by the Favé gauge during those current records against which the letter P is marked in the 'remarks' column in the tables. The Favé disks were examined under the microscope in the usual way, but in no case was there any measurable pressure fluctuation, other than the slow change due to the rise or fall of the tide. It may be concluded that the amplitude of any pressure fluctuation present must have been less than that due to 2 cm. of water.

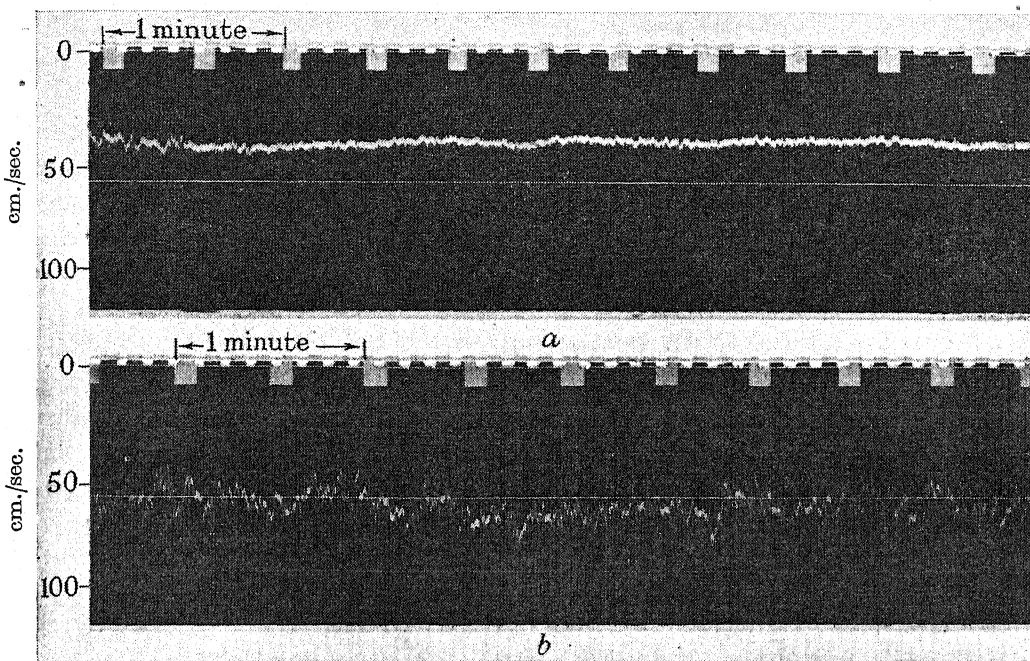


FIGURE 2. Sample lengths of two current records. *a*, no. 27, 1.8 m. below surface in a calm sea. *b*, no. 33, 1.1 m. above bottom, showing turbulent fluctuations.

5. DISCUSSION OF SHORT-PERIOD FLUCTUATIONS

5.1. *Effect of surface waves*

Since the short-period current fluctuations observed at Queen's Fort, and described in a previous paper, were attributed mainly to wave currents, we may consider, first of all, the influence of surface waves on the present results. The pressure amplitudes at the bottom were below the limit of sensitivity of the Favé gauge, i.e. less than 2 cm. of water. For waves of period 3 sec. the corresponding horizontal velocity would be 4.2 cm./sec. This is of the same order of magnitude as the values of u_1 given in tables 1 and 2, so that the evidence of the Favé records alone does not exclude the possibility of the fluctuations being due to wave currents.

There is reason to believe, however, that the wave pressures were considerably below the limit of sensitivity of the Favé gauge and the wave currents correspondingly

smaller. The waves at the surface did not exceed 2 ft. (60 cm.) in height during the series of observations and were usually smaller, while the period was not more than 3 sec. The depth of water was greater than 10 m. at Bromborough Buoy and greater than 17 m. off Rock Ferry. Waves of period 3 sec. and amplitude 30 cm. at the surface would be attenuated to give a pressure amplitude of 0.3 cm. of water at 10 m. and 0.02 cm. at 17 m. The corresponding amplitude of horizontal velocity would be 0.7 cm./sec. at 10 m. and 0.04 cm./sec. at 17 m. If the effective period were 4 sec. (and it was found previously that the apparent wave period increases with depth), the corresponding pressure amplitude would be 2.5 cm. at 10 m. and 0.6 cm. at 17 m., and the horizontal velocities 3.9 cm./sec. at 10 m. and 0.9 cm./sec. at 17 m. Unless the waves were travelling in a direction parallel to that of the current, the component of wave velocity in the direction of the mean current would be less than these figures, which refer, in any case, to extreme conditions.

It therefore seems improbable that wave currents could account for the amplitudes of the short-period current fluctuations recorded in tables 1 and 2, particularly in the deeper water position off Rock Ferry. Moreover, the general character of the records near the bottom differed from that of the records obtained at Queen's Fort, or of those in the present series obtained near the surface in the presence of waves.

5.2. *Dependence of fluctuations on mean current and height above the bottom*

Figure 3 shows the mean amplitude u_1 of the short-period fluctuations plotted against the mean current U for all records taken with the meter in the stand. There is an evident tendency for u_1 to increase with U , although the scatter of the points

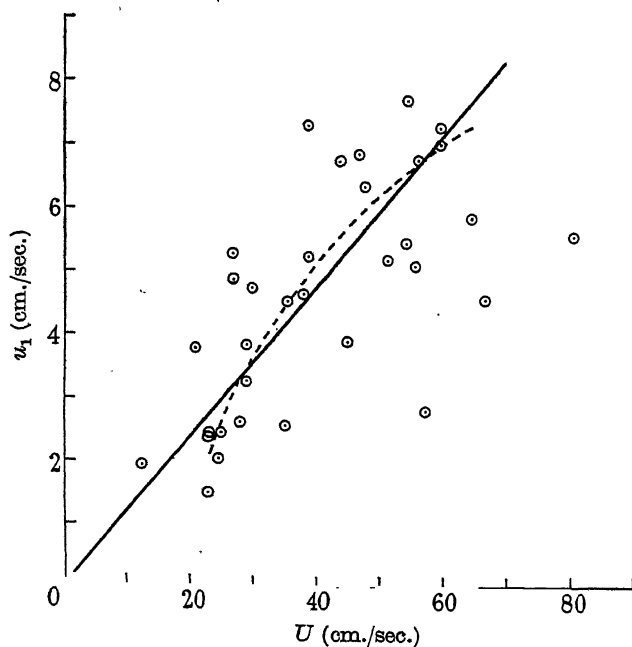


FIGURE 3. Short-period fluctuations near the bottom, showing dependence of mean amplitude of fluctuation u_1 on the mean current U .

is considerable. To a first approximation it seems justifiable to assume a linear relation between u_1 and U and take the ratio u_1/U as being independent of U . A rather better fit to the observations is given by the broken curve, suggesting that u_1 increases less rapidly for large values of U .

As shown in tables 1 and 2, records were obtained with the meter in three different positions in the stand, i.e. 0.6, 1.1 and 1.4 m. above the bottom. Mean values of u_1/U and its standard deviation are shown in table 4 for the three heights. The difference between the figures for 1.1 and 1.4 m. is insignificant, but the mean value of u_1/U for 0.6 m. is appreciably higher. The average for 0.6 m. is based on only six individual records, but, if statistical considerations may be applied, the probability of the difference between the mean for 0.6 m. and the combined mean for 1.1 and 1.4 m. being significant is over 90 %. In view of the small number of records at 0.6 m. and the wide scatter of the values of u_1/U for all three heights, no distinction has been made between the various heights in figure 3.

TABLE 4. DEPENDENCE OF FLUCTUATION RATIO ON HEIGHT ABOVE BOTTOM. (METER IN STAND)

height above bottom (m.)	short-period fluctuations		long-period fluctuations	
	mean u_1/U	standard deviation	mean u_2/U	standard deviation
0.6	0.144 ± 0.022	0.053	0.092 ± 0.015	0.037
1.1	0.113 ± 0.009	0.034	0.101 ± 0.012	0.044
1.4	0.109 ± 0.012	0.038	0.099 ± 0.009	0.029
all heights	0.117 ± 0.007	0.040	0.099 ± 0.007	0.037

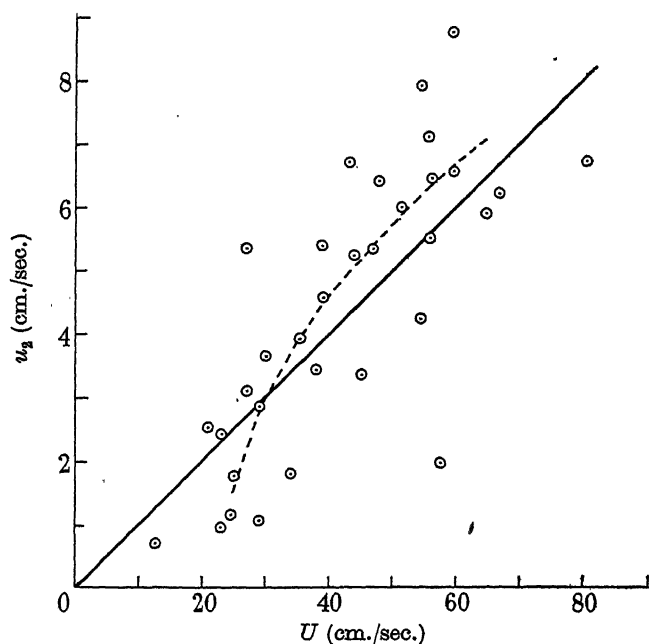


FIGURE 4. Long-period fluctuations near the bottom, showing dependence of mean amplitude of fluctuation u_2 on the mean current U .

Taking all the measurements near the bottom, the average value of u_1/U is 0.117, with a standard deviation of 0.040, and may be taken as applying to an average height of 1.1 m. above the bottom.

The values of the estimated mean period T_1 of the short-period fluctuations have been examined in a similar way, but there does not appear to be any significant correlation with the mean current or with the height above the bottom.

5.3. Dependence of fluctuations on depth below surface

Considering now the records taken with the meter suspended freely at various depths, figure 5*a* shows the values of u_1/U plotted against depth below surface for three series of observations. The difference between the first series (records 55 to 60) and the second series (records 61 to 66) is striking. In the first series the fluctuations are small near the surface and increase with depth, at first rapidly and then more slowly. In the second series the fluctuations are larger near the surface, decrease to a minimum at about 6 m. and then increase again at greater depths. At the time the first series was taken there was no wind and the sea was calm, while during the second series there was a wind of force 4 and surface waves of up to 2 ft. (60 cm.) in height. The character of record no. 61, at 3.0 m., was also different from the characters of records near the bottom, showing groups of regular, wave-like oscillations, as distinct from the irregular character of fluctuations near the bottom.

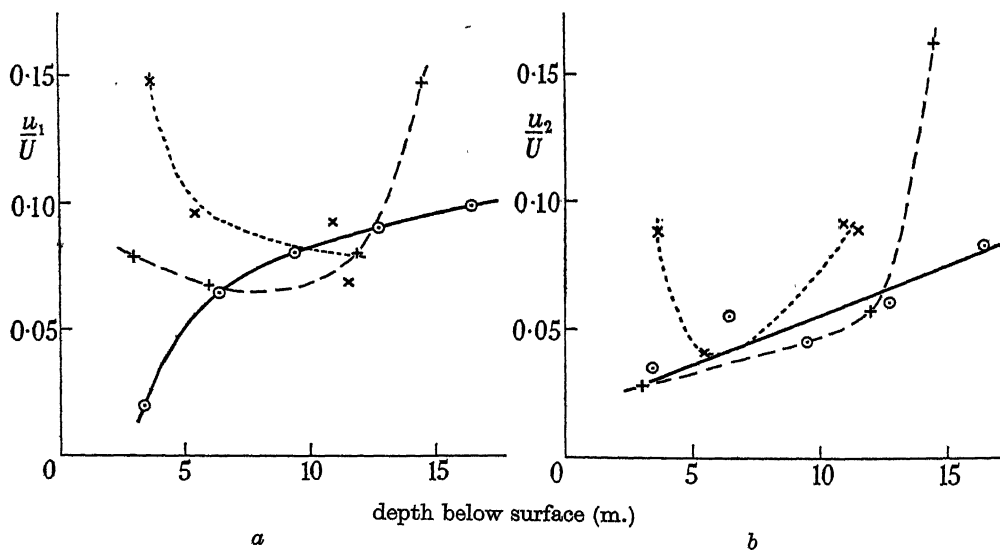


FIGURE 5. Dependence of fluctuation ratio on depth below surface, with the meter suspended freely. *a*, short and *b*, long-period fluctuations. $\circ$ — $\circ$ records 55 to 60; +—+ records 61 to 66; $\times$ — $\times$ records 67 to 70.

The third series of records (nos. 67 to 70) shows still larger fluctuations near the surface, decreasing with depth, but it does not seem possible to relate this to the state of the sea surface, which was intermediate between conditions for the first and second series. Record no. 69, at 3.7 m., was again different in character, showing

irregular fluctuations of large amplitude and mostly of shorter periods than usually occurred. In this respect it resembled records nos. 35 and 36 (near the bottom), which also showed large amplitude fluctuations of unusually short periods. Other records showed occasional bursts of such fluctuations which were not included in the mean values tabulated. These records have the common feature that they were all taken near times of high water, when there was considerable movement of shipping in the river. The times of passing of ships or tugs at distances of 50 to 300 yd. were noted on several occasions, and on examining the records the unusually short-period fluctuations were in evidence at approximately the times noted, although it was not possible to find any sudden onset.

5.4. Correlogram analysis

The auto-correlation of the short-period fluctuations as a function of time was examined from selected records. Enlarged tracings of the records were made and the trace measured at intervals of approximately 0.6 sec., smoothing out the oscillations of 1 sec. period which, as mentioned earlier, were due to the natural period of the galvanometer. It was also found necessary to eliminate the long-period fluctuations before computing the correlations, as otherwise the short-period variations were obscured. This elimination was done graphically. We have, in effect, selected fluctuations within a certain range of time scale, corresponding to periods between about 1 and 15 sec.

If u_t denotes the deviation of the velocity at time t from the long-period trend and $u_{t+\tau}$ that at time $t + \tau$, the auto-correlation coefficient for time interval τ is given by

$$r_\tau = \frac{\sum_{t=0}^{T-\tau} u_t u_{t+\tau}}{\left[\sum_{t=0}^{T-\tau} u_t^2 \sum_{t=0}^{T-\tau} u_{t+\tau}^2 \right]^{1/2}},$$

T being the total duration of the record analyzed. At first T was taken as 2 min., but it was found that the correlograms of successive 2 min. lengths of the same record differed so considerably that 2 min. did not seem to be long enough to contain a representative sample of the fluctuations. A length of 10 min., which appeared long enough to be free from this objection, was therefore adopted.

Figure 6 shows the correlograms of 10 min. lengths of four different records. Records nos. 32 and 33, the results of which are given in figure 6*a*, were both taken with the meter in the stand at a height of 1.1 m. above the bottom in a depth of water of 21 m. The values of the mean current U were 38 and 56 cm./sec. respectively. The two curves are very similar, r_τ falling to zero at $\tau = 3$ sec. approximately and then becoming negative, reaching a minimum of nearly -0.2 and then gradually increasing to attain small positive values after about 12 sec. Figure 6*b* shows the correlograms for two records, nos. 61 and 64, taken with the meter suspended freely at depths of 3.0 and 11.9 m. below the surface respectively. These curves each show oscillations whose amplitude decreases as τ increases superimposed on a curve of the type shown in figure 6*a*. The periods of the oscillations are approximately 3.4 sec. for the record at 3.0 m. depth and 4.6 sec. for that at 11.9 m. As surface waves of

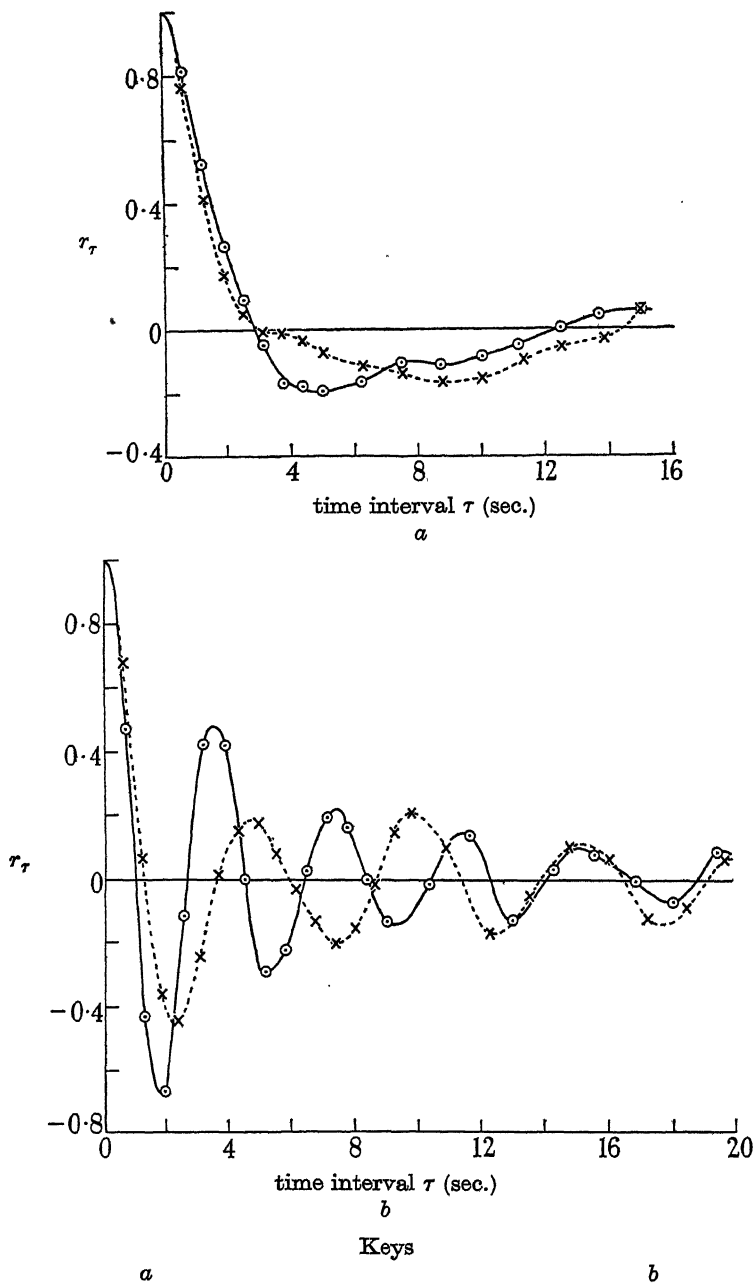


FIGURE 6. Auto-correlation curves for the short-period fluctuations. *a*, records near the bottom; *b*, records near the surface and at mid-depth, in the presence of surface waves.

period about 3 sec. and maximum height 60 cm. were observed at the time, it seems reasonable to attribute these oscillations to the effects of wave currents.

The general shape of the correlation curves in figure 6a is reminiscent of those obtained for the turbulent fluctuations of air speed in wind-tunnel experiments, although the time scale in our case is of the order of 100 times as long. Since a graphical method has been used to eliminate both the very short and the long periods, it does not seem justifiable to attempt a more quantitative interpretation of the correlation curves.

The results of the correlogram analysis received qualitative confirmation from spectra of the fluctuations very kindly obtained for us by the staff of the Oceanographic Section of the A.R.L., using the wave analyzer. Of several records examined, no. 61 (near the surface) showed predominant periods in the range 3 to 5 sec., while the records taken near the bottom indicated no predominant periods but a more or less uniform distribution of amplitudes throughout the spectral range covered by the analyzer.

While the correlogram analysis has shown the difference in character between the wave currents and the short-period turbulent fluctuations, it has confirmed the similarity in the time scales of the two phenomena. Whether there is any significance in this correspondence, even when the direct effects of surface waves are negligible, remains a problem for further investigation.

5.5. *Turbulent character of the fluctuations*

The above discussion indicates the following reasons for believing that the short-period fluctuations found in the present series of observations are to be identified with the turbulent fluctuations associated with bottom friction:

(1) The amplitude of the fluctuations is approximately proportional to the mean current.

(2) The amplitude of the fluctuations near the bottom is too large to be due to wave currents.

(3) Since the fluctuations were observed when the meter was in a stand on the bottom, rolling of the boat could have had no influence.

(4) In a calm sea, the fluctuations are small near the surface and increase with depth.

(5) The mean value of u_1/U , i.e. 0.117 at 1.1 m. above the bottom, is comparable with the estimated value of $[u^2]^{\frac{1}{2}}/U$, namely, 0.09, given in § 1. For harmonic fluctuations

$$u_1 = \sqrt{2} [u^2]^{\frac{1}{2}} \quad \text{and} \quad \sqrt{2} [u^2]^{\frac{1}{2}}/U = \sqrt{2} \times 0.09 = 0.127,$$

but such close agreement is fortuitous.

The scale of the disturbances is thus of an order of magnitude consistent with their being associated with bottom friction. For example, a period of 5 sec. in a mean current of 50 cm./sec. would correspond to the passage of eddies 2.5 m. in diameter.

(6) The correlogram analysis has shown that, unlike the wave currents observed nearer the surface, the fluctuations near the bottom have no predominant periodicity, but resemble the random variations usually associated with turbulence.

6. DISCUSSION OF THE LONG-PERIOD FLUCTUATIONS

The mean amplitudes u_2 of the long-period fluctuations have been analyzed in a way similar to those of the short-period fluctuations. Figure 4 shows that u_2 increases with the mean current and that the ratio u_2/U is approximately independent of U , while table 4 shows that there is no significant difference between the mean values of u_2/U at heights of 0.6, 1.1 and 1.4 m. above the bottom. The observations with the meter suspended freely, plotted in figure 5*b*, indicate that u_2/U increases with depth below the surface (except for the series of records nos. 67 to 70, where the reason suggested for the increase of short-period fluctuations, i.e. disturbance due to shipping, may possibly account also for the increase in long-period fluctuations near the surface).

Examination of the traces suggests that a wide range of periods were present. Frequently fluctuations of period 30 sec. to 1 min. were superimposed on fluctuations of period 4 to 5 min. Fluctuations of one period did not usually persist for more than a few oscillations, and different parts of the same 30 min. record often showed groups with different predominant periods, sometimes interspersed with intervals of negligible fluctuations. In this respect the 'mean periods' given in tables 1, 2 and 3 are misleading, since they were estimated by counting the number of peaks in a given time and thus depend mainly on the shorter periods present. A correlogram analysis was made of the long-period fluctuations on two selected records, nos. 43 and 47. The former showed a significant period of $5\frac{1}{2}$ min. on which fluctuations of period about 1 min. were superimposed, while the latter also showed periods of the order of 1 min. but no significant indication of longer periods.

These long-period fluctuations have periods in the same range as those found by Thorade, by Mosby, in the unpublished work in the *Zephyr*, referred to in § 1, and by us at Queen's Fort in Liverpool Bay. The present results confirm those found in the analysis of the Queen's Fort observations (Bowden 1947, § 4.3).

The dependence of the amplitude on the mean current and the lack of persistence of individual periods are consistent with the fluctuations being turbulent in origin. Periods of 1 to 5 min. in a mean current of 50 cm./sec., however, would correspond to the passage of eddies of 30 to 150 m. in diameter which, in a depth of 20 m., seem too large to be associated with bottom friction and shear motion in the current. This does not exclude the possibility of horizontal eddies set up in some other way.

An alternative explanation would be that they are due to internal waves. It is hoped to investigate at some future time the possible occurrence of internal waves in the Mersey, but it appears that internal waves with periods of several minutes might occur and that the horizontal velocities might be of the same order of magnitude as the observed fluctuations.

7. VERTICAL DISTRIBUTION OF TIDAL CURRENT

From all records taken with the meter in the stand, the amplitude of the M_2 constituent of the tidal current near the bottom was deduced to be 51.7 cm./sec. The number of records at other depths was insufficient to enable a study to be made of the variation of the current with depth, but the variation from the surface to

three-quarters of the total depth does not seem to have been large. Taking the average of all records with the meter suspended freely, the amplitude of the M_2 constituent was 90.4 cm./sec. Thus at 1 m. above the bottom the current was 0.57 of the average current throughout the upper three-quarters of the depth.

8. FURTHER OBSERVATIONS

We intend to continue the investigation of turbulence in a tidal current by making observations as follows:

(1) Measurements of the current in two positions simultaneously, by using two current meters separated (*a*) in the same vertical line, (*b*) in the line of the mean current, (*c*) in a horizontal line transverse to the mean current, both recording on the same photographic paper.

(2) More measurements of turbulent velocities at low values of the mean current, involving double anchoring of the boat to avoid the necessity for raising the stand every slack water.

(3) Measurements of the vertical distribution of the mean current.

REFERENCES

- Bowden, K. F. 1947 *Proc. Roy. Soc. A*, **192**, 403.
Doodson, A. T. 1940 *Hydrog. Rev.* **17** (1), 79.
Kármán, Th. v. 1935 *Proc. IV Internat. Congr. Appl. Mech.*, Cambridge, 1934, p. 63.
Mosby, H. 1946 *Bergens Mus. Aarb.*, Naturv. rekke, no. 3.
Thorade, H. 1931 *Rapp. Cons. Explor. Mer*, **76**, 59.
Thorade, H. 1934 *Ann. Hydrogr. Marit. Meteorol.* **62**, 365.

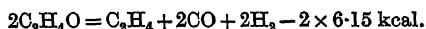
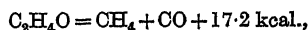
The ignition and flame reactions of ethylene oxide

BY F. A. BURDEN, PH.D. AND J. H. BURGOYNE, D.Sc., PH.D.

Department of Chemical Engineering and Applied Chemistry, Imperial College, London

(Communicated by Sir Alfred Egerton, F.R.S.—Received 11 December 1948)

Spontaneous ignition pressure-temperature relationships of ethylene oxide-oxygen and -air mixtures have been studied. A low-temperature ignition system exists and cool flames can be *initiated* in air at atmospheric pressure and below. The propagation of *cool* flames through cold mixtures under this condition has not, however, been observed. Nevertheless, ethylene oxide itself propagates a *decomposition* flame and is inflammable at constant volume in admixture with air for all percentages above 3.6. The products of decomposition are almost entirely accounted for by the two overall reactions:



These reactions are also very significant in the non-luminous flames that have been found to be propagated through the richer mixtures of ethylene oxide with air. It is shown that the latter flames are initiated only at high temperatures, of the same order as for normal flames, and that they correspond with the non-luminous flames occurring near to the high-temperature spontaneous ignition limit of richer mixtures.

INTRODUCTION

In recent years ethylene oxide has assumed some commercial importance as an organic intermediate, but prior to this its chief use was as a fumigant for the control of insect pests in foodstuffs. It has undoubtedly a number of properties that recommend it for the latter purpose, but its inflammability is a serious disadvantage, and on several occasions ethylene oxide explosions have occurred during fumigations.

Previous investigators have considered the question of reducing the inflammability of ethylene oxide in air. Early studies of its physiological properties had shown that its pesticidal effect is promoted by the presence of carbon dioxide, which increases the respiration rate of the insect. For this reason it has for some years been the custom to introduce 10 % of carbon dioxide into commercial ethylene oxide for use as a fumigant, a practice which also raises the vapour pressure to a convenient level for discharge from the cylinder. It is not surprising therefore that carbon dioxide was early considered for the purpose of suppressing its inflammability.

Jones & Kennedy (1930) completely determined the effect of carbon dioxide on the limits of inflammability of ethylene oxide in air, using the standard apparatus and technique adopted at the U.S. Bureau of Mines (Coward & Jones 1938). Their results were confirmed by R. M. Jones (1933), and later Peters & Ganter (1938), although using different experimental conditions, obtained a partly similar result. The three sets of observations are plotted together in figure 1.

On comparing the inflammability curves of Jones & Kennedy and of R. M. Jones in figure 1 with extinction limit curves of the normal type, one is immediately struck by the unusual 'bulge' on the rich-mixture side of the diagram. The seemingly most

obvious explanation of this peculiarity is that in this part of the field flames of the 'cool' type are involved—a suggestion that has already been advanced by Coward & Jones (1938) and which receives some support from the description given by the authors of the appearance of the rich-mixture flames. The analogy with the results of White (1927) and of Townend & Chamberlain (1937) on the effect of reduced pressure on the inflammability of diethyl ether is clear, and fairly recently a very similar composite dilution curve has been obtained by Jones & Scott (1947) for the system dimethyl ether-difluorodichloromethane-air. The only difficulty about this explanation arises from the fact that both Jones & Kennedy and R. M. Jones used a flame source throughout their tests for igniting the ethylene oxide mixtures. Where cool flames are involved one would hardly expect this source to be effective—a source of lower temperature would be needed. On the other hand, Peters & Ganter, who used an electric spark source, failed to obtain the bulge (figure 1). This result is consistent with the 'cool-flame' explanation. It is therefore quite uncertain what type of flame is really propagated in the rich mixtures.

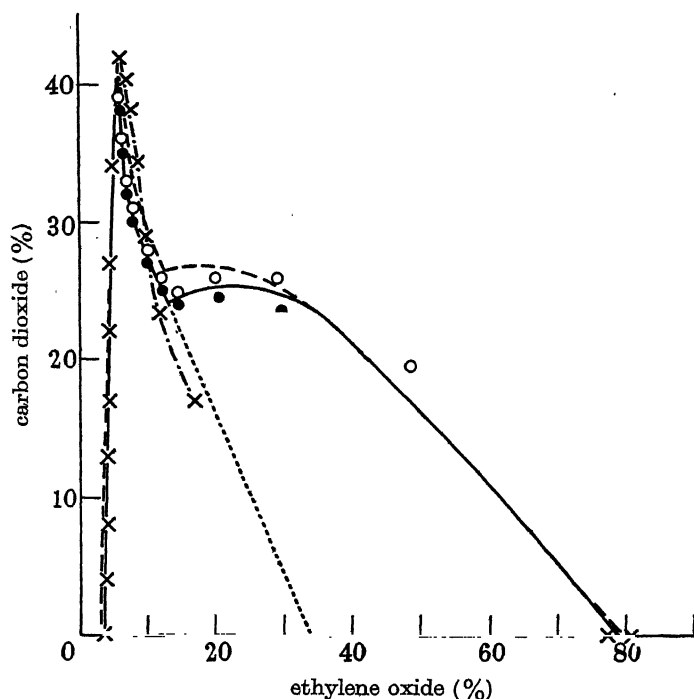


FIGURE 1. Results of determinations of the influence of carbon dioxide on the limits of inflammability of ethylene oxide in air. — G. W. Jones & Kennedy. $\times$ — $\times$ Peters & Ganter. — flame propagated $\bullet$: failed $\circ$ R. M. Jones. extrapolation after Coward & Jones. R. M. Jones's points have been omitted on the lower limit side of the curve as they substantially coincide with it.

The practical importance of the issue is of course that if ethylene oxide can propagate cool flames in air at atmospheric pressure, the minimum ignition temperatures are likely to be very low, perhaps as low as in the case of diethyl ether (c. 180°C).

If this were so, hitherto unsuspected sources might be responsible for ethylene oxide explosions, for the cool flame of a rich mixture could lead to normal ignition if it travelled into a region where the mixture was more dilute.

The investigation into the inflammability of ethylene oxide was therefore begun with two immediate objects in view:

(1) To see whether ethylene oxide-air or -oxygen mixtures have a low-temperature ignition system, and if so, to ascertain the minimum temperature and pressure conditions for the initiation of cool flames.

(2) To investigate the character of the peculiar rich-mixture flames observed by earlier workers and the conditions for their initiation and propagation.

Although the investigation arose from the practical considerations indicated, it led to results of interest with regard to combustion theory.

EXPERIMENTAL

Spontaneous ignition limits

The first objective was sought using the 'hot-vessel' method to determine ignition pressure-temperature limits for ethylene oxide-air or -oxygen mixtures. Preliminary experiments were made with a 5% ethylene oxide-air mixture in steel pressure apparatus similar to that of Townend & Mandlekar (1933). The results are given in figure 2, in which the ignition limit is shown as a function of temperature and pressure.

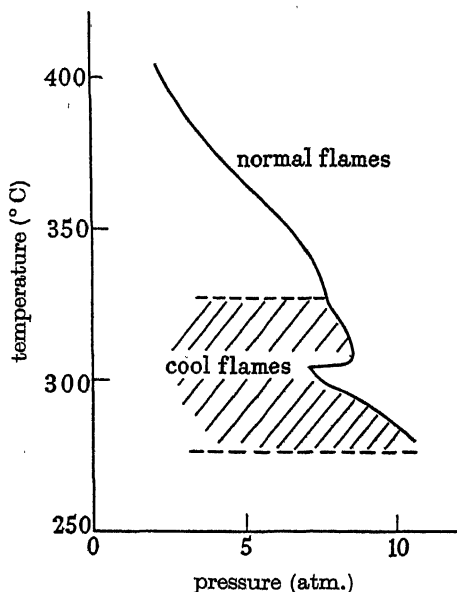


FIGURE 2. Ignition temperature-pressure limits of 5% ethylene oxide-air mixture in steel explosion vessel.

A low-temperature ignition peninsula occurs in the range 280 to 310°C. At lower pressure, in the range 275 to 325°C, very mild pressure pulses were observed on admitting the mixture to the heated vessel. There was little doubt that these were

due to cool flames, but owing to the brevity of the lag before the flame and the feebleness of the pulse at lower pressures, the cool flame pressure-temperature limit could not be defined with any certainty.

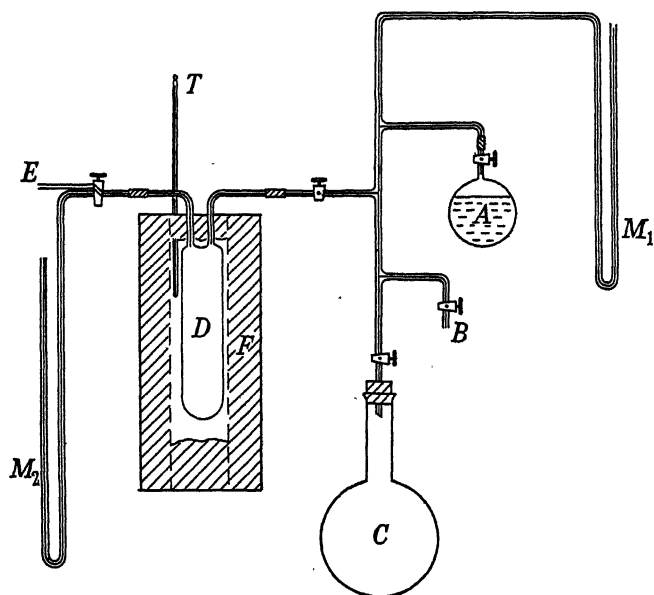


FIGURE 3. Low-pressure apparatus for determination of ignition limits. *A*, ethylene oxide storage bulb; *B*, to drying towers and O_2 storage; *C*, mixture storage bulb; *D*, reaction vessel; *E*, to vacuum pump; *F*, electric furnace; M_1, M_2 , manometers; *T*, thermometer.

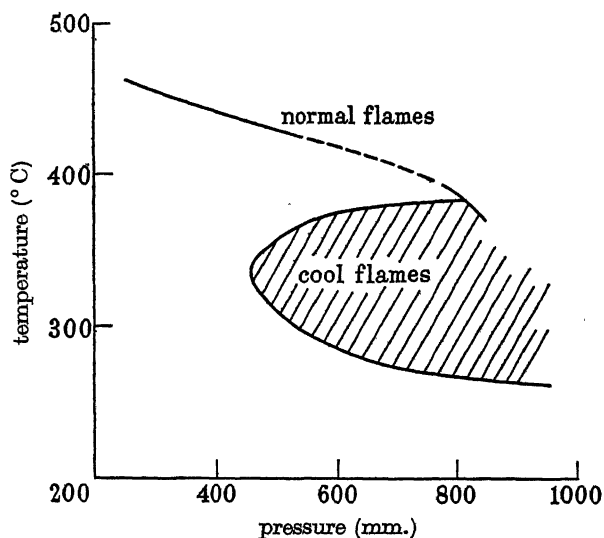


FIGURE 4. Ignition temperature-pressure limits of equimolecular ethylene oxide-oxygen mixture in quartz explosion vessel.

In consequence, the investigation was pursued using ethylene oxide-oxygen mixtures in a low-pressure apparatus, the details of which are shown in figure 3. The reaction vessel, which in this case was a quartz cylinder 5 cm. in diameter and

of about 550 ml. capacity, was heated to any desired temperature in an electric furnace. A stock of the mixture for test was made up by pressure in the 2l. globe. From this stock a part could be admitted rapidly to the thoroughly evacuated explosion vessel to give some previously decided initial pressure. The occurrence of ignition or otherwise was observed with the aid of the mercury manometer.

The ignition limits thus determined for an equimolecular ethylene oxide-oxygen mixture are shown in figure 7. At low temperatures, ignitions occurred at too high a pressure and with too great violence to allow the limit to be plotted in this apparatus. The ignition limit at high temperatures and the cool-flame limit could, however, be satisfactorily defined, although in the latter case some difficulty was at first experienced owing to the short lags and the feebleness of the cool-flame pulse near the limit. Nowhere in the cool-flame region was more than one pulse observed, in contrast with results for aldehydes (Newitt, Baxt & Kelkar 1939). Measurements of limits as in figure 4 were reproducible provided the explosion vessel was first 'standardized' with a few high-temperature ignitions after each cleaning.

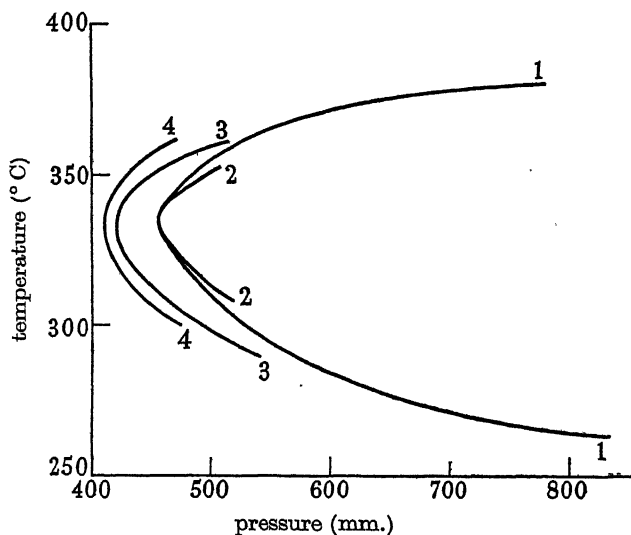


FIGURE 5. Effect of ethylene oxide-oxygen mixture composition on the cool flame pressure-temperature limit. Curves: 1, $(\text{CH}_2)_2\text{O} + \text{O}_2$; 2, $(\text{CH}_2)_2\text{O} + 2\text{O}_2$; 3, $2(\text{CH}_2)_2\text{O} + \text{O}_2$; 4, $4(\text{CH}_2)_2\text{O} + \text{O}_2$.

Figure 5 shows the results of measurements of the cool-flame pressure-temperature limit using mixtures containing ethylene oxide and oxygen in different proportions. While a reduction in the ethylene oxide: oxygen ratio from 1:1 to 1:2 had little effect, increasing the proportion of ethylene oxide lowered the minimum pressure for cool-flame initiation quite markedly. The optimum temperature for cool flames was, however, much the same for all mixtures, and any lowering effect that there may have been on the cool flame-true ignition pressure limit was insufficient to allow it to be determined in the low-pressure apparatus.

Figure 6 shows the effect of the presence of nitrogen on the cool flame temperature-pressure limit of a 2:1 ethylene oxide-oxygen mixture. Once more the optimum

temperature for the cool flames is not materially altered, but dilution lowers the minimum partial pressure of the reactants necessary for the initiation of cool flames. This conclusion suggests that some stage of the reactions leading to cool-

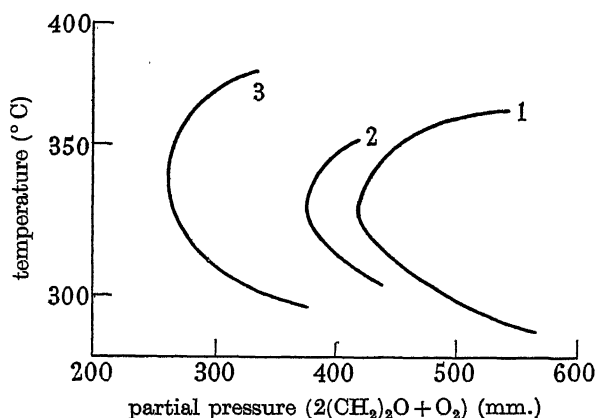


FIGURE 6. Effect of nitrogen dilution on the cool flame (partial) pressure-temperature limit for a 2:1 ethylene oxide-oxygen mixture. Curves: 1, $2(\text{CH}_2)_2\text{O} + \text{O}_2$; 2, $2(\text{CH}_2)_2\text{O} + \text{O}_2 + \frac{1}{2}\text{N}_2$; 3, $2(\text{CH}_2)_2\text{O} + \text{O}_2 + 4\text{N}_2$.

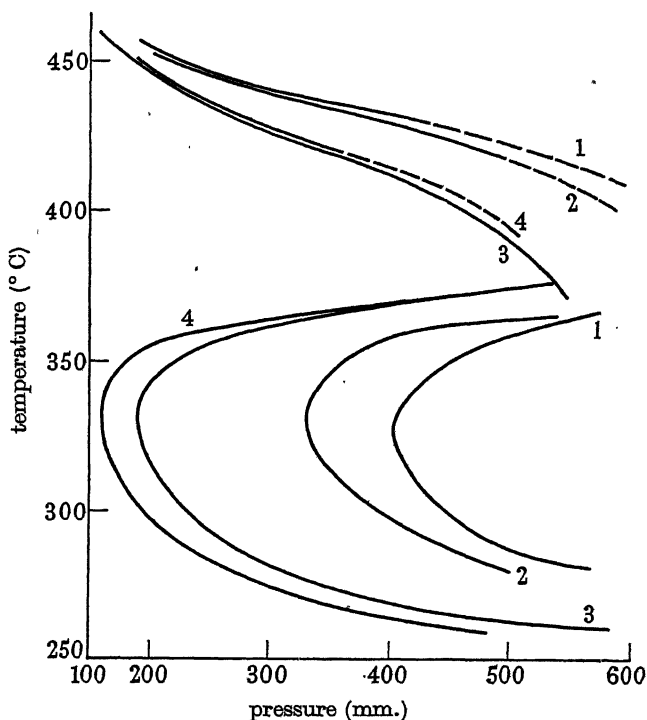


FIGURE 7. Effect of the addition of small amounts of acetaldehyde on pressure-temperature limits for an equimolecular ethylene oxide-oxygen mixture. Curves: 1, $(\text{CH}_2)_2\text{O} + \text{O}_2$; 2, $(\text{CH}_2)_2\text{O} + \text{O}_2 + 0.1\% \text{CH}_3\text{CHO}$; 3, $(\text{CH}_2)_2\text{O} + \text{O}_2 + 1\% \text{CH}_3\text{CHO}$; 4, $(\text{CH}_2)_2\text{O} + \text{O}_2 + 2\% \text{CH}_3\text{CHO}$.

flame initiation (not necessarily the cool-flame reaction itself) is a chain reaction in which the active centres suffer deactivation at the vessel wall.

The mixture referred to in figure 6 which contains 1 vol. of 2:1 ethylene oxide-oxygen mixture with $\frac{4}{3}$ vol. of nitrogen is equivalent to a 28.6 % ethylene oxide-air mixture. The results show that it is possible to *initiate* a cool flame in ethylene oxide-air mixtures well below atmospheric pressure (at a minimum of 610 mm. in this particular case). Whether the flame can be *propagated* through such mixtures at atmospheric temperature and pressure is a question to which we shall presently revert.

Figure 7 shows the observed effects of small additions of acetaldehyde upon the pressure-temperature limits for high-temperature ignition and cool flames in an equimolecular ethylene oxide-oxygen mixture. The high-temperature ignition limit was quite definitely lowered in these circumstances, but the cool-flame pressure limit was more strikingly affected. The presence of 0.1, 1 and 2 % of acetaldehyde lowered the minimum pressure for cool-flame initiation by 20, 52 and 60 % respectively. At the same time, the optimum temperature for cool flames was not greatly altered. Enough observations were made to show that the cool flame-true ignition pressure limit was sufficiently lowered in the presence of these small amounts of acetaldehyde to bring much of it within the range of the apparatus, but unfortunately the ignitions obtained were so violent as to damage the explosion vessel and experiments were discontinued.

Limits of inflammability

Among other significant results, the work outlined in the preceding section showed that cool flames can be initiated in ethylene oxide-air mixtures at and below atmospheric pressure. The first of the two objects mentioned in the introduction was therefore attained with such a result as to leave the question implied in the second still open. It was clear that the minimum pressure for cool-flame propagation through a cold mixture must be greater than that for its initiation at high temperature, but by how much was uncertain. Moreover, in view of the observed lowering effect of diluent (nitrogen) on the cool-flame pressure limit it was just conceivable that the rich-mixture flame area in the presence of carbon dioxide as found by Jones & Kennedy and R. M. Jones (figure 1) might even yet refer to a cool flame. The question was therefore investigated directly, at atmospheric and reduced pressures.

For this purpose the apparatus illustrated in figure 8 was employed. It consisted of a 5 cm. diameter Pyrex glass tube about 130 cm. long supported in a light framework which allowed the tube to be swung about its central transverse axis and brought to rest at any required angle for a flame-propagation test. The ends of the tube were carefully squared and smoothed and were closed by oversized rubber corks pressed against them. One of these carried the igniting source and an admission tube for gases, and the other was adjustably spring-loaded to provide excess pressure relief.

The mixture for test was made up in the tube by partial pressures and was mixed by rotating the tube and framework so that a light metal bobbin fell from end to end of the tube and back again during each complete rotation. It was found that twenty-

five complete journeys by the bobbin ensured adequate mixing. After this any slight adjustment (reduction) of the total pressure of the mixture was made, the tube set at the desired elevation, the bobbin supported magnetically at the end remote from the ignition source and the mixture fired with the room in darkness.

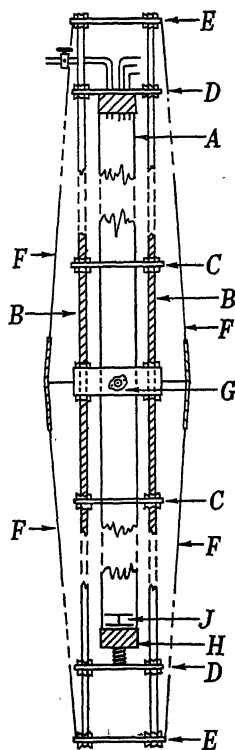


FIGURE 8. Apparatus for the measurement of limits of inflammability at atmospheric and reduced pressures. *A*, Pyrex tube; *BB*, threaded rods; *CC*, plywood supports; *DD*, plywood end-supports; *EE*, metal end-plates; *FF*, bracing wires; *G*, centre block and rotation bearing; *H*, spring-loaded rubber cork; *J*, mixing bobbin.

At first, both magneto-spark and platinum-filament ignition were employed. At atmospheric pressure they gave very similar results, but at low pressure the platinum filament was more effective, and it was therefore employed for the systematic measurement of inflammability limits. The results shown in figure 9 refer to a vertical tube with the source at the lower end. They reveal a number of interesting features.

At atmospheric pressure, flame was propagated throughout the range 3.6 to 100 % of ethylene oxide. Up to about 68 % changes in the flames were normal, the non-luminous lower limit flame gradually giving place to the luminous upper limit flame. At this point, however, the luminous flame became changed, with increasing proportions of ethylene oxide, into a pale blue flame with a luminous tail, from which the luminosity had entirely disappeared at about 70 %. This type of flame persisted up to 100 %, and in pure ethylene oxide a still paler but more rapid decomposition flame was propagated.

The effect of reduced pressure on the boundary compositions for these flames is shown in figure 9. The lower limit of inflammability was very little affected by reduced pressure, but the upper (luminous flame) limit was rapidly reduced (broken line). Decomposition flames of pure ethylene oxide were propagated down to 633 mm., but thereafter an upper composition limit for the blue flames appeared which fell rapidly with diminishing pressure. The normal and blue flame areas had individual pressure minima at 78 and 72 mm. (almost flat) respectively.

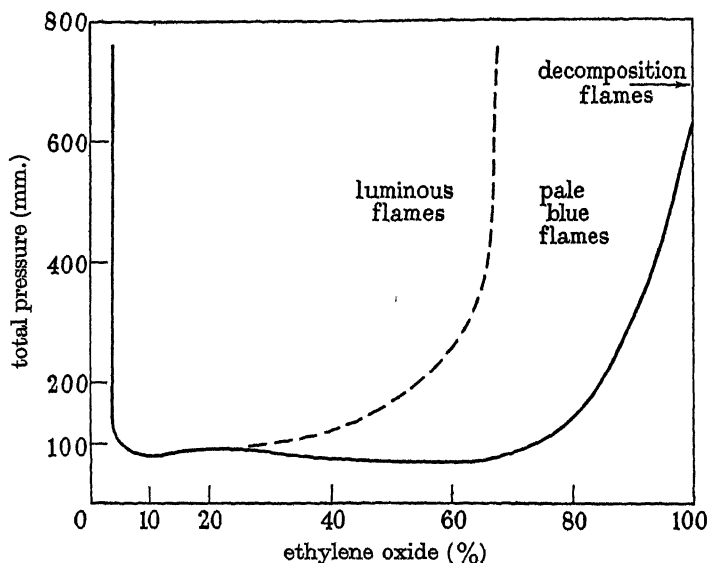


FIGURE 9. Effect of pressure (reduced) on inflammability ranges of ethylene oxide in air (closed tube; propagation vertically upward).

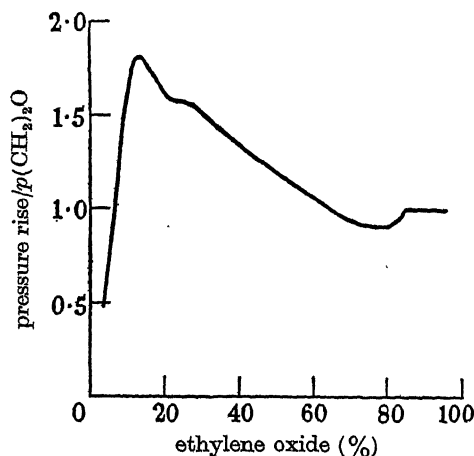


FIGURE 10. Relationship between ratio permanent pressure increase/initial pressure ethylene oxide and mixture composition for flame propagation at the pressure limit.

At the same time that limit determinations were made, the pressure increases (after cooling) due to inflammation of the mixtures at the pressure limit were noted.

In figure 10 the ratio pressure increase: initial pressure of ethylene oxide ($\Delta p/p$) is plotted against mixture composition. Commencing with the lower limit mixture at atmospheric pressure, $\Delta p/p$ rises steeply to a peak at 13 to 14 % ethylene oxide and then falls continuously to 80 %, apart from an inflexion between 20 and 30 % corresponding with the low maximum in the pressure-limit curve (figure 9). Above 85 to 90 % measurements were not valid, since the pressure during the actual propagation was sufficient to cause the excess pressure relief to operate.

The determinations of pressure limits were repeated with downward propagation in the tube held vertically. In these circumstances the blue and decomposition flames were not propagated, and limits almost the same as those for the normal flame in figure 9 were obtained.

Although, in its failure to propagate downwards, the blue flame (which evidently corresponds with that observed by Jones & Kennedy and R. M. Jones) partook of the character of a cool flame, its light emission and flame speed appeared somewhat greater than experience with diethyl ether and acetaldehyde would have led us to expect. It seemed, moreover, that it could be *propagated* through cold mixtures at a much lower pressure than that at which the cool flame could be even *initiated* in a hot vessel. In order therefore to determine conclusively the nature of the blue flame, the temperature required for its initiation by a hot source was compared with that required to initiate the normal flames. For this purpose the platinum filament already in use could not be employed because ignition was always preceded by catalytic combustion on the filament which caused it to self-heat to an extent that was beyond control. After some preliminary experiments a 'non-catalytic' igniter was developed which consisted of about 5 cm. of 36 s.w.g. nichrome wire, with copper leads, completely sealed into a drawn-out Pyrex tube and formed into a helix about 5 mm. in diameter. Although fragile this igniter was satisfactory, giving surface temperatures up to about 750° C. It was fitted into the end of the flame propagation tube in place of the platinum filament and connected in series with a variable resistance and sensitive ammeter to the source of electrical supply. The igniter was calibrated for maximum temperature versus current by means of a very fine platinum-platinum/rhodium thermocouple placed in contact with its hottest part, care being taken to have inert gas in the tube at the approximate pressure used for mixture tests during the calibration.

When this igniter was used to ignite ethylene oxide-air mixtures there was no evidence of surface combustion. Using the current through the igniter to measure temperature, the relationship was obtained for typical ethylene oxide-air mixtures between pressure and the surface temperature of the igniter required to cause ignition.

The results are shown in figure 11. The mixtures chosen were:

(1) 16.7 % (approximately equimolecular mixture ethylene oxide-oxygen), propagating a normal flame.

(2) 44.4 % (ethylene oxide:oxygen = approximately 4:1), propagating a blue flame at low pressure and a normal luminous flame at atmospheric pressure.

(3) 75 % giving a blue flame at all pressures up to atmospheric.

In addition, the ignition of pure ethylene oxide was studied.

The minimum temperature of the igniter necessary to effect ignition always fell to a constant figure with increasing pressure and the pressure reached a constant value (representing the corresponding limit in figure 9) for high source temperatures. For pure ethylene oxide the bend in the curve was very abrupt. Limiting source temperatures observed were as follows:

% mixture	°C
16.7	640
44.4	639
75	650
100	675

Although a temperature gradient must exist between a local igniting source and the adjacent gas mixture and therefore the above figures are not in any sense actual ignition temperatures, it is perfectly clear from them that the blue and normal flames, and, indeed, the decomposition flames too, are initiated at temperatures of similar orders. The blue flames therefore cannot possibly be cool flames.

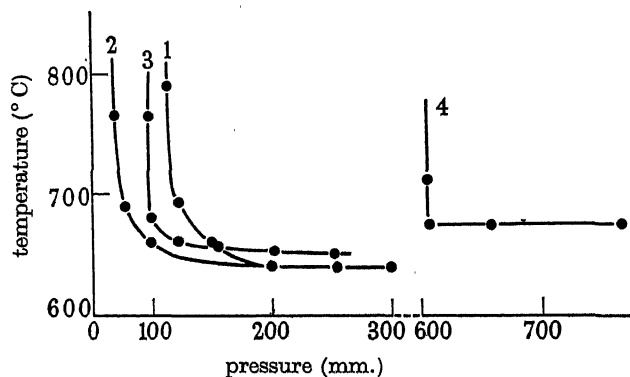


FIGURE 11. Relationship between pressure and source temperature for local ignition of ethylene oxide-air mixtures and pure ethylene oxide. Percentages of $(\text{CH}_2)_2\text{O}$ in curves 1 = 16.7 %, 2 = 44.4 %, 3 = 75 %, 4 = 100 %.

Comparative study of normal, blue and decomposition flames

It being established that the blue flame was of a high-temperature origin, we recalled observations made by one of us in connexion with the ignition of cyclohexane (Burgoyne, Tang & Newitt 1940), and by Newitt & Thornes (1937) on the ignition of propane. In these instances it had been noticed that at the high-temperature spontaneous ignition limit there was a zone of pressure within which blue flames were formed, accompanied by a relatively mild pressure pulse and some subsequent slow reaction. At higher pressure these gave place to yellow or orange ignitions in which reaction was complete in one stage and accompanied by normally vigorous pressure pulses. It was necessary therefore to ascertain whether the blue flames propagated through cold ethylene oxide-air mixtures were in fact related to phenomena occurring near to the spontaneous ignition limit at high temperatures.

The spontaneous ignition apparatus (figure 3) was modified to enable the explosion vessel to be observed visually through a tube projecting up the furnace from below

and closed at the lower end by a prism which allowed the vessel to be viewed horizontally. Pressure-temperature limits were then determined at high temperatures for a number of ethylene oxide-air mixtures and also for the decomposition of ethylene oxide itself. At the same time many visual observations of the ignitions were made.

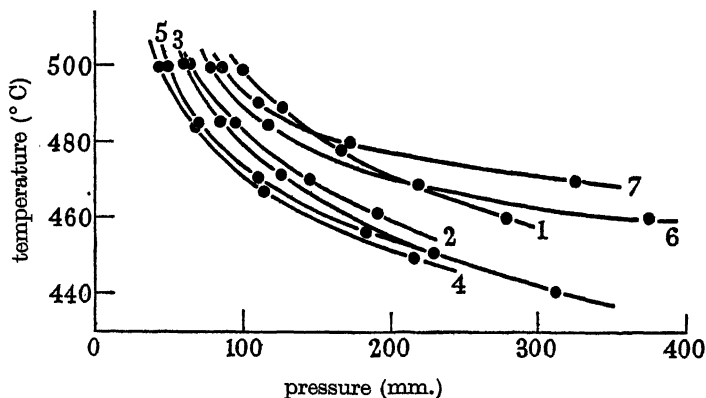


FIGURE 12. Spontaneous ignition temperature-pressure relationships for ethylene oxide and some mixtures with air (high temperature range). Percentages of $(\text{CH}_2)_2\text{O}$ in curves 1 = 16.7 %, 2 = 30 %, 3 = 44.4 %, 4 = 75 %, 5 = 85 %, 6 = 95 % and 7 = 100 %.

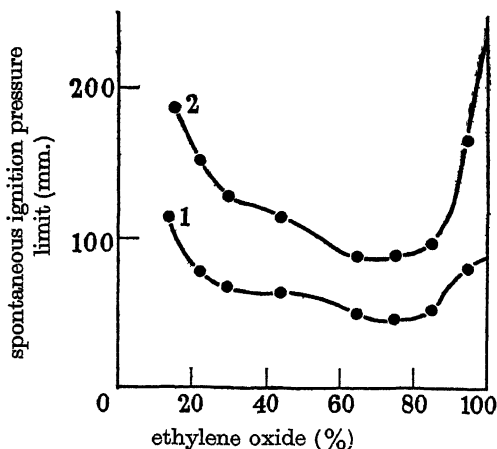


FIGURE 13. Relationship between spontaneous ignition pressure and mixture composition for ethylene oxide-air mixtures at 500° C (curve 1) and 475° C (curve 2).

As a result it was found, as had been thought possible, that over a certain pressure range above the ignition limit the ethylene oxide-air flames were blue, becoming yellow above this pressure range. The width of the blue area was small at lower and greater at higher temperatures, and at a given temperature it increased with increasing proportion of ethylene oxide in the mixture. The decomposition flame of ethylene oxide was in all cases very faint and only just visible in complete darkness.

In view of these observations, the ignition pressure-temperature limit curves, shown in figure 12, were employed to derive the relationship between spontaneous

TABLE 1. PRODUCTS OF THE SPONTANEOUS IGNITION OF ETHYLENE OXIDE AND OF ITS MIXTURES WITH AIR

initial % C_2H_4O in air	temp. (° C)	initial pressure (mm.)	moles per 100 moles C_2H_4O		type of flame	final pressure (mm.)	products: moles per 100 moles C_2H_4O					
			O_2	N_2			CO_2	O_2	C_2H_4	CO	H_2	CH_4
16.7	470	355	104.3	395.8	yellow	482	25.2	0.7	0	183.8	93.0	13.7
44.4	470	195	26.2	99.2	blue	317	4.2	0.7	15.7	116.6	78.6	32.2
75	470	307	7.0	26.4	blue	505	0.4	1.5	15.0	85.2	50.4	35.4
75	470	244	7.0	26.4	blue	—	2.2	0	15.1	86.8	57.0	36.2
100	510	360	0	0	decomp.	720	0.7	0.2	21.7	99.8	60.5	56.0
100	620	34	0	0	decomp.	68	1.1	0.4	21.8	101.8	40.0	53.6

TABLE 2. PRODUCTS OF FLAMES PROPAGATED THROUGH ETHYLENE OXIDE VAPOUR AND ITS MIXTURES WITH AIR AT 18° C

nominal % C_2H_4O in air	initial pres- sure (mm.)	moles per 100 moles C_2H_4O		type of flame	final pres- sure (mm.)	products: moles per 100 moles C_2H_4O							residual C_2H_4O
		O_2	N_2			CO_2	O_2	C_2H_4	CO	H_2	CH_4	CH_3CHO	
16.7	98	104.3	395.8	yellow	126	28.8	4.2	7.2	176.2	100.0	17.4	—	—
44.4	105	26.2	99.2	blue	165	5.1	3.4	18.8	103.2	46.5	41.2	faint	6.0
75	121	7.0	26.4	blue	204	4.8	1.4	16.7	83.6	32.2	45.6	trace	16.6
100	760	0	0	decomp.	*	1.5	0	11.6	72.8	19.8	50.2	2.7	23.4

* Blew off at pressure release.

ignition pressure and mixture composition at specified temperatures. The results for 500 and 475° C are given in figure 13. The similarity to figure 9 was at once apparent and suggested that it was probably correct to relate the behaviour on the high-temperature spontaneous ignition limit with that in the flame propagation tube. To obtain more conclusive evidence, however, comparative chemical analyses were made on the products formed in flame reactions under the two sets of conditions.

For this purpose a simple sampling arrangement was constructed which could be inserted in the vacuum lead of either apparatus (i.e. figures 3 and 8). It consisted of a Pyrex spiral trap with a small bulb at the lower end and fitted with stopcocks to inlet and outlet. The outlet was connected via a T-piece and other stopcocks to a gas burette and the vacuum pump. In taking product samples, the trap was cooled in an alcohol-solid carbon dioxide mixture and evacuated. Some of the product was then drawn through the trap by operating the gas burette and the rest withdrawn by the vacuum pump, a sample of the uncondensed gas being retained in the burette.

When examining the products in the flame-propagation tube, each experiment was done twice, one lot of condensed product being used to determine residual ethylene oxide and the other for aldehydes, etc. Analysis of the two gas samples provided a check on reproducibility. With the much smaller spontaneous ignition apparatus explosions were repeated several times under the same conditions to get sufficient material for examination. Methods of chemical analysis are detailed in the appendix. The condensable products collected were of course the total formed, but the gaseous products were related to reactants through the nitrogen content.

Products were analyzed from 16.7, 44.4 and 75 % ethylene oxide-air mixtures and from pure ethylene oxide, reacting in each case by spontaneous ignition at high temperature and by the passage of flame through the cold mixture. In every instance a pressure a little above the limit for the flame reaction was chosen. The detailed conditions and results are set out in tables 1 and 2.

The principal products were gaseous, the condensable part from flame propagation reactions containing only unchanged ethylene oxide, small amounts of acetaldehyde, traces of formaldehyde, and water (not estimated directly). No acids of any kind were detected at any time. From the spontaneous ignition reactions there was neither residual ethylene oxide nor aldehyde of any kind, but a little tarry matter was formed from the spontaneous decomposition of ethylene oxide. In the flame-propagation reaction of ethylene oxide the pressure increase was such as to cause some unburnt mixture to blow off through the pressure relief. For this reason a true balance could not be obtained in this case.

DISCUSSION

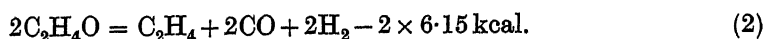
Explosive decomposition

A superficial consideration of the experimental results described leads to the conclusion that the behaviour of ethylene oxide-air or -oxygen mixtures on inflammation must be markedly influenced by the decomposition reaction of ethylene oxide itself, especially where the richer mixtures are concerned. It will be shown, in fact, that the peculiar behaviour of the rich mixtures is largely accounted for by

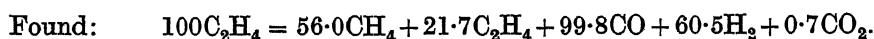
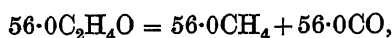
the decomposition reaction. For this reason it is appropriate first to consider what has been learned about this reaction.

The slow decomposition of ethylene oxide vapour has been studied in some detail by previous workers (Heckert & Mack 1929; Thompson & Meissner 1936; Seddon & Travers 1936), but the explosive reaction has received no attention, although it seems likely that Peytral (1926) did get ignitions when she passed the vapour through a platinum tube at high temperature.

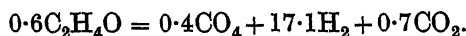
The products of the decomposition reactions shown in tables 1 and 2 may be most simply explained by assuming the occurrence of simultaneous reactions having the following overall effects:



Thus for the spontaneous decomposition at 510° C (table 1) we have:

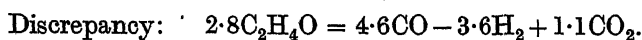
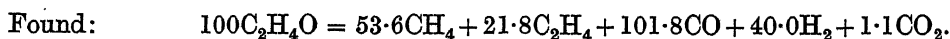
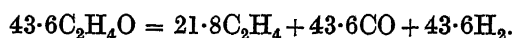
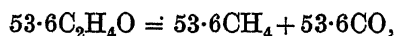


The discrepancy is:

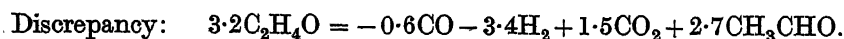
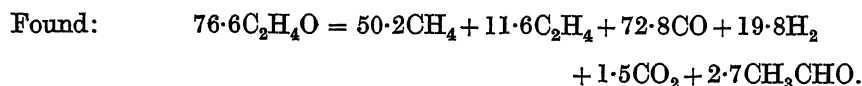
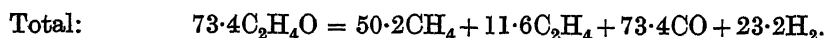
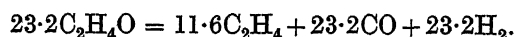
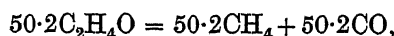


The large discrepancy for H_2 corresponds with a discrepancy in the hydrogen balance and is thought to be due to an error on the estimation of free hydrogen in the gas sample.

Again, for the spontaneous decomposition at 620° C (table 1):



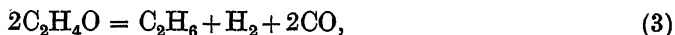
And for the flame propagation through a cold mixture:



Allowing for the isomerization of ethylene oxide the discrepancy is



The results could equally well be interpreted by replacing reaction (1) by similar ones involving larger paraffin molecules, e.g.



but there is no evidence from our work of such reactions and they are certainly less probable than (1).

With regard to reaction (2), on the other hand, our results do not seem to permit of any alternative interpretation, for to assume that the unsaturated product (measured by absorption) is an olefine higher than ethylene would imply an increase in the ratio of CO/unsaturateds over that actually found.

Most workers on the slow decomposition of ethylene oxide have agreed that reaction (1) is prominently involved under their experimental conditions, and Seddon & Travers (1936) found evidence of (2) as a significant parallel reaction. Heckert & Mack (1929) and Thompson & Meissner (1936) suggested that (3) occurs to a small extent, but did not mention (2).

It will be noticed that the overall reactions expressed by equations (1) and (2) are respectively exothermic (17.2 kcal.) and endothermic (12.3 kcal.). The occurrence of ignition must therefore depend on a preponderance of (1) over (2) as regards thermal effects. In table 3 are given the relative proportions of reactions (1) and (2) that occurred in the three cases already analyzed and hence the corresponding nett evolution of heat.

TABLE 3. NETT HEAT EVOLUTION IN THE EXPLOSIVE
DECOMPOSITION OF ETHYLENE OXIDE

conditions of reaction	temp. (° C)	pressure (mm.)	ratio of reactions (1):(2)	nett evolution heat, kcal./mole C ₂ H ₄ O
spontaneous ignition	620	34	1.23	7.0
spontaneous ignition	480	360	1.29	6.7
propagation	18	760	2.19	10.3*

* Includes an allowance of 0.5 for $\text{C}_2\text{H}_4\text{O} = \text{CH}_3\text{CHO} + 23.1$ kcal.

An evolution of 9.7 kcal./g.mole from reactions (1) and (2) occurring in the ratio 2.19:1 (propagation reaction, table 3) leads to a calculated flame temperature increase of about 500° C for the decomposition flame. Although this figure is low, it is well above the known temperatures of combustion 'cool flames' (maximum about 400° C).

Attention has been drawn by the previous investigators (Heckert & Mack 1929; Thompson & Meissner 1936) to the fact that the kinetics of the slow decomposition are not uniform. Thus there is a change of the order of reaction from unimolecular at pressures above 400 mm. to bimolecular at lower pressures (Thompson & Meissner 1936). At the same time the activation energy decreases slightly (Thompson &

Meissner 1936). Our observations upon the ignition limit are in conformity with these results. Thus in figure 14 is shown the relationship between $\log_{10} p/T$ and $1/T$ for the ignition limit of ethylene oxide (curve 7) derived from figure 12, curve 7. The relationship is not linear as it would be if the kinetics of the reaction underlying ignition were uniform with changing temperature or pressure, and the direction of curvature is such as to imply an increase of order and/or a decrease in activation energy with diminishing pressure or increasing temperature.

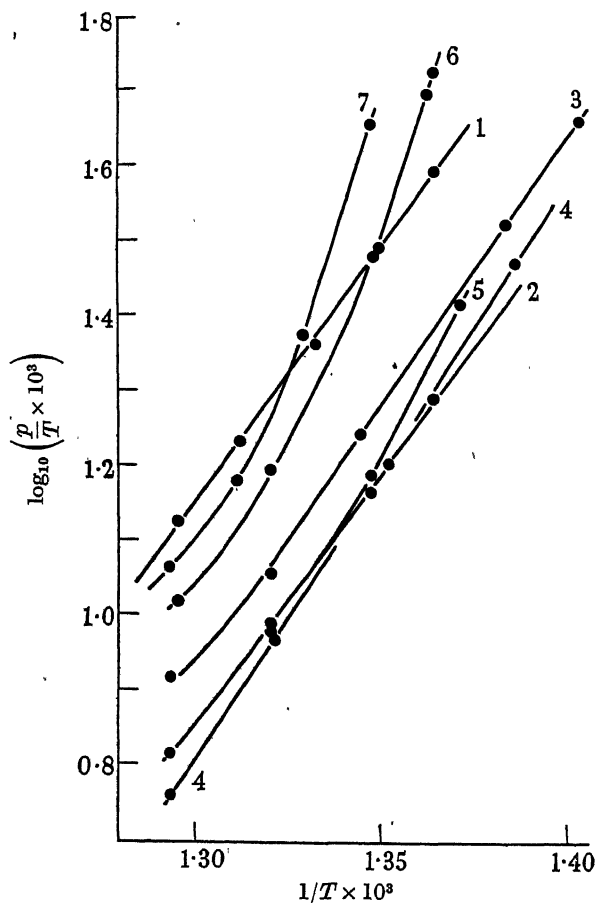


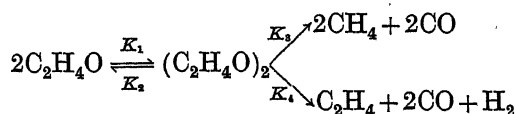
FIGURE 14. $\log p/T$ against $1/T$ curves for the ignition limit of ethylene oxide and its mixtures with air. Curves 1 = 16.7 %, 2 = 30 %, 3 = 44.4 %, 4 = 75 %, 5 = 85 %, 6 = 95 % $\text{C}_2\text{H}_4\text{O}$ in air. Curve 7 = pure $\text{C}_2\text{H}_4\text{O}$.

Thompson & Meissner interpret the change of order as an instance of quasi-unimolecular behaviour. They find, by equating the observed number of molecules reacting at the pressure p where the unimolecular constant begins to fall off with the product of the collision number and the precise Maxwell energy distribution factor for n degrees of freedom, that n has the very reasonable value of 8.

It has been shown elsewhere (Burgoyne & Burden 1949) that the presence of nitrogen has no lowering effect upon the ignition pressure of ethylene oxide. This

supports the opinion of Thompson & Meissner and others that the decomposition reactions are not of a chain character.

It appears then that the decomposition of ethylene oxide proceeds by a quasi-unimolecular collision reaction having the total effect expressed by equations (1) and (2) taken in certain proportions. This result may be accounted for by supposing that a sufficiently energetic collision between two molecules of ethylene oxide results in the formation of a short-lived complex which may either revert to the original two molecules due to collision with a third molecule of ethylene oxide, or may decompose by either of two paths, yielding respectively the ultimate products according to equations (1) and (2). This scheme may be expressed thus:



where K_1 , K_2 , K_3 and K_4 are velocity constants for the steps indicated (which in the last two cases will probably be composite). In these circumstances the intermediate $(\text{C}_2\text{H}_4\text{O})_2$ will assume a stationary concentration (A) and we have

$$K_1(\text{C}_2\text{H}_4\text{O})^2 - K_2(A)(\text{C}_2\text{H}_4\text{O}) = (rK_3 + (1-r)K_4)(A),$$

where $r(1-r)$ is the ratio in which steps 3 and 4 occur. Hence

$$(A) = \frac{K_1(\text{C}_2\text{H}_4\text{O})^2}{K_2(\text{C}_2\text{H}_4\text{O}) + rK_3 + (1-r)K_4},$$

and the overall rate of reaction

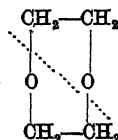
$$(rK_3 + (1-r)K_4)(A) = \frac{K_1(\text{C}_2\text{H}_4\text{O})^2}{\frac{K_2}{rK_3 + (1-r)K_4}(\text{C}_2\text{H}_4\text{O}) + 1},$$

which tends to $\frac{K_1(rK_3 + (1-r)K_4)}{K_2} \times (\text{C}_2\text{H}_4\text{O})$ (unimolecular) when the pressure is high, and to $K_1(\text{C}_2\text{H}_4\text{O})^2$ (bimolecular) when the pressure is low. If K_3 and K_4 differ somewhat and r varies under different conditions the kinetics will be complicated, but otherwise the occurrence of parallel reactions will not greatly affect the issue.

The structure of the intermediate $(\text{C}_2\text{H}_4\text{O})_2$ is somewhat speculative, but if, for

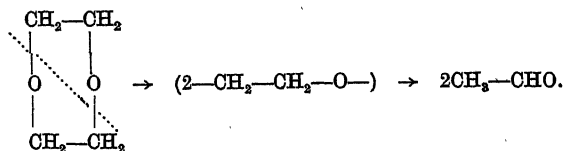
the purpose of illustration, we assume $\begin{array}{c} \text{CH}_2-\text{CH}_2 \\ | \quad | \\ \text{O} \quad \text{O} \\ | \quad | \\ \text{CH}_2-\text{CH}_2 \end{array}$, three reasonable possibilities present themselves regarding its fate:

(1) Collision with another molecule may lead to the scission

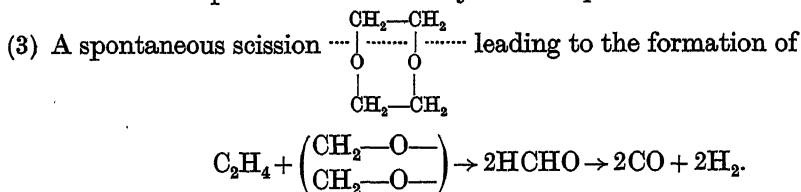


re-formation of ethylene oxide.

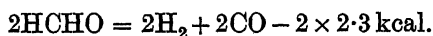
(2) A spontaneous scission of the same type leading to the exothermic formation of acetaldehyde ($C_2H_4O = CH_3CHO + 23.1 \text{ kcal.}$)



The observed activation energy of the ethylene oxide decomposition together with the exothermal effect of the isomerization would be sufficient to initiate the known endothermic decomposition of acetaldehyde to $CH_4 + CO$.



The thermal effects of these steps are:



Combustion reaction at high temperatures

It has already been suggested that the decomposition flame propagated through cold ethylene oxide corresponds rather closely with the flame initiated at the same pressure somewhat above the limiting ignition temperature, i.e. with less lag than at the limit. A similar comparison can be made for the high-temperature flames of ethylene oxide burning in air. It has been shown, however, that in the latter case there is a zone just within the spontaneous ignition limit in which pale blue flames are formed, and which becomes wider with (a) decreasing pressure (increasing temperature) and (b) increasing proportion of ethylene oxide in the mixture. Thus the likelihood that the flame propagated through a mixture will be a blue flame is increased with (a) decreasing pressure, or (b) increasing richness of the mixture. Qualitatively, this describes the form of the boundary in figure 9 between the luminous and blue flames. Further evidence of correspondence is provided by comparing the shapes of pressure-composition limits for spontaneous ignition (figure 13) and for flame propagation (figure 9), and by comparing the products of reaction of similar mixtures under the two conditions (tables 1 and 2).

Certain observations already recounted point to the existence of two rather distinct high-temperature flame reactions between ethylene oxide and air, one in rich mixtures and one in those of around stoichiometric composition. These observations can be enumerated briefly:

(1) The change-over from luminous to pale blue flames in both spontaneous ignition and propagation experiments.

(2) The inflexion in the pressure limit curves between 20 and 30 % ethylene oxide (figures 9 and 13).

(3) The inflexion in the $\Delta p/p$ curve between 20 and 30 % ethylene oxide (figure 10).

To these should be added the evidence of figure 14 which shows the $\log p/T - 1/T$ relationship at the spontaneous ignition limit for various ethylene oxide-air mixtures as well as for ethylene oxide itself. Those graphs relating to the 16.7, 30, 44.4 and 75 % mixtures are essentially linear (the first two showing no curvature whatever), and adhere to the relationship $\log_{10} p/T = A/T + B$. If the ignition limit were thermal in origin the gradient of these lines (A) would be equal to $E/2.30 \times 2 \times R$. Thus values for E , the activation energy, may be derived as in table 4.

TABLE 4. ACTIVATION ENERGIES FROM SPONTANEOUS IGNITION
LIMITS OF ETHYLENE OXIDE-AIR MIXTURES

% C ₂ H ₄ O	A	$E = 2.30 \times 2 \times 1.98 \times A$ (kcal./mole)
16.7	6.80	61.9
30	6.73	61.4
44.4	7.12	64.8
75	7.67	69.9

The graphs for the 85 and 95 % mixtures show an increasing resemblance to the curve for pure ethylene oxide, already discussed, and undoubtedly reveal the increasing part played by the decomposition reaction in the inflammation of the richer mixtures. The analytical results shed further light on this aspect.

In tables 5 and 6 the results set out in tables 1 and 2 are further analyzed. Here are shown first of all the extent to which the products of the several reactions may be directly attributed to the decomposition reactions (1) and (2) as judged respectively by the amounts of methane and ethylene surviving. The products found but not accounted for directly by the decomposition reactions are set out under the heading 'differences'. These products must be attributed to oxidation reactions either of ethylene oxide itself, of its isomer acetaldehyde which can be formed from it, or of the products of the decomposition reactions (1) and (2). The oxidation of the hydrocarbons from the decomposition reactions would leave corresponding amounts of carbon monoxide and hydrogen and this might account for part of the amounts of these products surviving.

It is obvious from tables 5 and 6 that oxidation reactions play a much greater part in the formation of the yellow flame with the 16.7 % mixtures than in any other case. Since the oxidation reactions are much more exothermic than the decompositions, this leads to a much higher flame temperature. Although the results are insufficient for plotting, it appears that the survival of the direct decomposition products is much greater with the 44.4 % mixtures than might be expected from a comparison of results for the 75 and 16.7 % mixtures. In this there is a further indication of the 'inflection' in the characteristics of the reaction as the luminous flame gives place to the non-luminous.

Combustion reactions at low temperatures

It has been shown that cool flames are initiated in ethylene oxide-oxygen mixtures below atmospheric pressure and in the temperature range 270 to 380° C. This range is roughly that which is characteristic of the higher hydrocarbons. For acetaldehyde the temperature range is somewhat lower (Newitt, Baxt & Kelkar 1939).

TABLE 5. PRODUCTS OF THE SPONTANEOUS IGNITION OF ETHYLENE OXIDE AND OF ITS MIXTURES WITH AIR

initial temp. (° C)	initial % C ₂ H ₄ O	initial pres- sure (mm.)	moles O ₂ per 100 moles C ₂ H ₄ O	flame appearance	products: moles per 100 moles C ₂ H ₄ O						C ₂ H ₄ O not accounted by (1) and (2)
					reaction (1)		reaction (2)		differences		
					CH ₄	CO	C ₂ H ₄	CO	H ₂	CO ₂	
470	16.7	355	104.3	yellow	13.7	13.7	0	0	170.1	93.0	86.2
470	44.4	195	26.2	blue	32.2	32.2	15.7	31.4	53.0	47.2	36.4
470	75	307	7.0	blue	35.4	35.4	15.0	30.0	19.8	20.4	34.6
470	75	244	7.0	blue	36.2	36.2	15.1	30.2	20.4	26.8	33.6
510	100	360	0	decomp.	56.0	56.0	21.7	43.4	0.4	17.1	0.6
620	100	34	0	decomp.	53.6	53.6	21.8	43.6	4.6	-3.6	2.8

TABLE 6. PRODUCTS OF FLAMES PROPAGATED THROUGH ETHYLENE OXIDE VAPOUR AND ITS MIXTURES

nominal % C_2H_4O	initial pres- sure (mm.)	moles O_3 per 100 moles C_2H_4O	flame appearance	products: moles per 100 moles C_2H_4O						residual C_2H_4O (3)	differences			C_2H_4O not accounted by (1), (2) and (3)
				reaction (1)		reaction (2)		CO	C_2H_4		H_2	CO	H_2	
				CH_4	CO									
16.7	98	104.3	yellow	17.4	17.4	7.2	14.4	14.4	0	144.4	85.6	28.8	68.2	
44.4	105	26.2	blue	41.2	41.2	18.8	37.6	37.6	6.0	24.4	8.9	5.1	15.2	
75	121	7.0	blue	45.6	45.6	16.7	33.4	33.4	16.6	4.6	-1.2	4.8	4.4	
100	760	0	decomp.	50.2	50.2	11.6	23.2	23.3	26.1*	-0.6	-3.4	1.5	0.4	

* Including 2.7 moles isomerized to CH_3CHO .

The reactions leading to the formation of the cool flame of ethylene oxide are evidently accelerated by the presence of nitrogen, suggesting that chain reactions subject to wall deactivation are involved. Acetaldehyde in small quantities also exerts a promoting effect, suggesting that this material alone or in conjunction with oxygen is capable of providing the necessary chain carriers. Moreover, the presence of acetaldehyde lowers quite markedly the minimum temperature for the initiation of cool flames. Thus it seems that the chain carriers that initiate (at a lower temperature) the cool-flame process of acetaldehyde itself are the same that (at higher temperatures) initiate the ethylene oxide cool flame. This observation is consistent with the view (although it does not establish it) that ethylene oxide must first isomerize to acetaldehyde to give the chain carriers necessary to initiate the cool flame, and that it is this process which determines the minimum temperature for cool flames in this instance. It is difficult, however, to see how chain carriers could be such as to influence the isomerization of ethylene oxide, and hence it may well be that the process which leads to the initiation of cool flames with ethylene oxide does not actually involve the isomerization reaction.

The two-stage ignition has not been realized in the quartz apparatus owing to pressure limitations, but experiments in the high-pressure apparatus show that it can occur. A few indications were obtained with the quartz apparatus that the pressure limit for such ignitions is lowered in the presence of acetaldehyde, probably for just the same reason that the reactions leading to cool-flame initiation are promoted by the same additive.

General conclusions

Experiments on spontaneous ignition have shown that cool flames can be *initiated* in ethylene oxide-air mixtures at atmospheric pressure. Efforts to cause them to propagate through a cold mixture at this pressure have, however, been unsuccessful, and under such conditions only high-temperature flames have been propagated. However, in a closed vessel all mixtures with air containing more than 3.6 % ethylene oxide, as well as ethylene oxide itself, have been found to be inflammable by a high-temperature source. Above about 70 % of ethylene oxide the flame is non-luminous and invisible in daylight. It is clear that the peculiar 'bulge' in the extinction limit curves of previous workers is due not to cool flames but to the high-temperature 'blue' flame which in turn is closely associated with the explosive decomposition of ethylene oxide itself.

The influence of constant pressure conditions on the limits is among the matters of interest that it is proposed to elucidate in the extension of this work, and it may be that therein will be discovered the reason why Jones and others found an upper limit for ethylene oxide in air. In the $\Delta p/p$ composition curve for the flame-propagation limit (figure 10) there is a slight indication of transition at 80 % which is the upper limit previously found.

A further investigation into the kinetics of the slow-combustion and cool-flame reactions of ethylene oxide is now in hand. In view of this, detailed suggestions regarding the combustion mechanism have not been made in the present communication.

APPENDIX

Specification of materials

Ethylene oxide was received in cylinders containing 10 % carbon dioxide which was entirely removed by drawing off the liquid and distilling under reduced pressure. Only the middle one-third of the distillate was retained for use. It contained a trace of acetaldehyde.

Oxygen and nitrogen were derived from commercial cylinders.

Acetaldehyde was redistilled from the 'pure' material purchased and the fraction distilling at 21° C retained.

Analytical methods

Gases were analyzed in a Bone-Newitt apparatus using 10 % potash for CO₂, alkaline pyrogallol for oxygen, mercuric sulphate for ethylene and ammoniacal cuprous chloride for CO. Hydrogen and hydrocarbons were determined by explosion in excess of oxygen and measurement of the contraction and CO₂ formation. Residual oxygen was removed, followed by treatment with ammoniacal cuprous chloride, after which the remaining gas was assumed to be nitrogen.

Products condensed at -78° C. Separate samples were taken from duplicate experiments for determination of (a) ethylene oxide, (b) other condensable products. For ethylene oxide, 25 ml. of 0.1 N-hydrochloric acid in saturated magnesium chloride solution was sucked into the trap and left for 30 min. The solution was then poured into a conical flask. Solution and washings were titrated separately against standard caustic soda using phenol red as indicator and the total residual acid found. A blank determination was carried out under similar conditions.

For aldehydes, the contents of the trap were dissolved in water and made up to 50 ml. If a qualitative test with Schiff's reagent gave a positive result, total aldehydes were estimated in an aliquot part of the solution by Ripper's bisulphite-iodine method. Formaldehyde was estimated colorimetrically by a method based on the qualitative test described by Steigmann (1942). A reagent consisting of 0.025 % rosaniline acetate, 2.5 % vol./vol. sulphuric acid and 0.05 % thioglycollic acid was added to an aliquot part of the solution, and the brownish purple colour compared with standards prepared from formaldehyde in known concentrations. The difference between total aldehyde and formaldehyde was taken to be acetaldehyde, but its presence was demonstrated qualitatively by adding 1 ml. of 10 % alcoholic piperidine solution followed by a few drops of freshly prepared 10 % sodium nitroprusside solution and observing development of the characteristic blue colour (or greenish, with traces).

The need for this investigation was stressed in an unpublished report by one of us (J.H.B.) with Mr G. Williams-Leir, made to the Fire Research Division, Research and Experiments Department, Home Office, in 1945. Reference to this report is made with the permission of the Rt Hon. Viscount Falmouth who was the Head of the Fire Research Division at that time. Some of the information referred to in the introduction to the present paper, especially that summarized in figure 1, was collected by Mr Williams-Leir.

A suggestion regarding the need for the investigation also came to us independently through Professor J. W. Munro, to whom we are indebted for help in the matter of apparatus and materials.

The gas analyses were carried out for us by Mr J. Hooper to whom our thanks are due.

Finally, one of us (F.A.B.) is indebted to the Gas Light and Coke Company Ltd. for their Fellowship, during the tenure of which the work was carried out.

REFERENCES

- Burgoyne, J. H. & Burden, F. A. 1949 *Nature*, **163**, 723.
Burgoyne, J. H., Tang, T. L. & Newitt, D. M. 1940 *Proc. Roy. Soc. A*, **174**, 379.
Coward, H. F. & Jones, G. W. 1938 *Bull. U.S. Bur. Min.* no. 279 (revised).
Heckert, W. W. & Mack, E. 1929 *J. Amer. Chem. Soc.* **51**, 2706.
Jones, R. M. 1933 *Industr. Engng Chem.* **25**, 394.
Jones, G. W. & Kennedy, R. E. 1930 *Industr. Engng Chem.* **22**, 146.
Jones, G. W. & Scott, F. E. 1947 *Rep. Invest. U.S. Bur. Min.* no. 4125.
Newitt, D. M., Baxt, L. M. & Kelkar, V. V. 1939 *J. Chem. Soc.* p. 1703.
Newitt, D. M. & Thornes, L. S. 1937 *J. Chem. Soc.* p. 1656.
Peters, G. & Ganter, W. 1938 *Angew. Chem.* **51**, 29.
Peytral, E. 1926 *Bull. Soc. chim.* (iv), **39**, 206.
Seddon, R. V. & Travers, M. W. 1936 *Proc. Roy. Soc. A*, **156**, 234.
Steigmann, A. 1942 *J. Soc. Chem. Ind.* **61**, 18.
Thompson, H. W. & Meissner, M. 1936 *Trans. Faraday Soc.* **32**, 1451.
Townend, D. T. A. & Chamberlain, E. A. C. 1937 *Proc. Roy. Soc. A*, **158**, 415.
Townend, D. T. A. & Mandlekar, M. R. 1933 *Proc. Roy. Soc. A*, **141**, 484.
White, A. G. 1927 *J. Chem. Soc.* p. 498.

The production of mesons by electrons

By I. N. SNEDDON AND B. F. TOUSCHEK, *Department of Natural Philosophy,
University of Glasgow*

(Communicated by N. F. Mott, F.R.S.—Received 20 January 1949—Revised
27 April 1949)

The cross-section of the production of mesons in collisions between electrons and nuclei is calculated. It is assumed that the electrons interact with the nucleons through the electromagnetic field. The results are given for the case of pseudoscalar mesons. It is found that for electrons of energy 300 MeV the cross-section is 2×10^{-30} cm.² per nucleon.

1. INTRODUCTION

With the development of high-energy accelerators of the synchrotron type the production of mesons in the laboratory has become a practical possibility. Already mesons have been produced by the heavy particles accelerated in the Berkeley synchro-cyclotron (1948). The advanced state of the construction of several synchrotrons designed to accelerate electrons to energies in the 300 MeV region makes it desirable to consider theoretically the production of mesons in collisions between electrons and atomic nuclei.

In the case of the synchrotron we are faced with the possibility of using either electrons or γ -rays as primary particles. The production of mesons by γ -rays has been considered recently by Morette & Peng (1947), who find for 300 MeV γ -radiation a cross-section of the order of 10^{-27} cm.² per nucleon. It should, however, be observed that this value is rather high as a result of the authors' assumption of a rather low rest mass for the π -meson. The experimental study of the production of mesons by means of γ -rays produced in the target of a synchrotron suffers from the disadvantage that, instead of a mono-energetic source, a continuous spectrum of γ -rays has to be employed. Because of the well-known difficulties of γ -ray spectroscopy in this energy region it is almost impossible to determine experimentally the variation with energy of the cross-section of the production of mesons by γ -rays. In the case of γ -rays this dependence on the energy is the most important prediction that can be made theoretically.

On the other hand, if we use electrons as bombarding particles a certain disadvantage arises from the fact that the interaction between electrons and nuclei is less, by a factor of the order of the fine-structure constant, than that between γ -rays and nuclei. Indeed, as a consequence of the Weizsäcker-Williams method (1934) in which the electron is replaced by a spectrum of equivalent photons, we expect that every process arising from the interaction between γ -rays and nuclei has its counterpart with electrons but, owing to the low intensity of the equivalent spectrum, the corresponding cross-section is lower.

Certain advantages do, however, follow from the use of electrons. In the first place the beam of electrons from a synchrotron is approximately mono-energetic, which will permit an experimental test to be made to ascertain whether or not the

dependence of the cross-section on the energy of the electron beam is in agreement with theoretical predictions. Secondly, since in the final state of the system there are three particles (the inelastically scattered electron, the nucleus and the liberated meson), a continuous meson spectrum will result. From the theory we can calculate the distribution of energy of the mesons produced by electrons of a given energy. This calculation therefore yields a second relationship which is capable of being compared directly with experimental evidence.

In this paper we shall consider the inelastic scattering of electrons in an energy range of 200 to 1000 MeV leading to the production of π -mesons, i.e. those mesons which are thought to be responsible for nucleon-nucleon interactions. If we denote the charge of the target nucleus by Z we may write

$$e + (Z) \rightarrow e' + (Z \mp 1) + \pi^{\pm}, \quad (1)$$

for the process under consideration. The method used for the calculation of the cross-section for this process is based upon Born's approximation, but an estimate of the effect of the electrostatic field of the nucleus on the emitted meson is given. We consider, in particular, the pseudoscalar meson theory developed by Møller & Rosenfeld (1940), Chang (1942) and others. The upper limit of 1000 MeV to the validity of the results of this paper is set by the fact that we neglect the recoil of the nucleus during the collision process. This recoil is closely related to the retardation phenomena associated with the electromagnetic interaction between the electrons and the protons of the nucleus. The lower limit of 200 MeV (for a meson of rest energy 150 MeV, say) is set by the condition that the kinetic energy of the meson should be large in comparison with the binding energy of a nucleon in the nucleus. This has its origin in the fact that in the calculation we treat the various possible excited states of the nucleus as being energetically equivalent.

Within these limits the dependence of the cross-section on the energy of the incident electron, the mass and charge of the nucleus and that of the emitted meson is given by the formula (24) below. The factor $g(E)$, depending on the energy E of the incident electron, has the approximate forms (28) and (29) in the non-relativistic and extreme-relativistic regions respectively. For electrons of energy 300 MeV the cross-section varies roughly as A (the mass number of the target nucleus), ranging from 2×10^{-30} cm.² for a single nucleon to approximately 10^{-28} cm.² for the heaviest nuclei.

The distributions of energies of positive and negative mesons emitted from a very heavy nucleus are shown in figure 1 for the case of 300 MeV electrons and a meson mass of 150 MeV. For any other energy in the range of validity of the formulae (200 to 1000 MeV roughly) the energy spectrum is given by equation (21) below.

2. CALCULATION OF THE DIFFERENTIAL CROSS-SECTION

We wish to calculate the differential cross-section for the production of a pseudo-scalar meson with momentum q by an electron of momentum p and energy $E \simeq p$. This last relation is true in the system of units in which the units of mass, velocity and action are respectively the electronic mass m , the velocity of light c and Planck's

constant $\hbar$. In this system of units e^2 —the square of the electronic charge—is $4\pi/137$.

The differential cross-section is determined by the following method. We first calculate the electromagnetic potentials $(\mathbf{p}' | A_\nu | \mathbf{p})$ corresponding to a transition of the electron from a state with momentum $\mathbf{p}$ to one with momentum $\mathbf{p}' = \mathbf{p} - \mathbf{k}$. If the current corresponding to the transition of the nuclear system from a state with 'bound' or no mesons to a state with a free meson of momentum $\mathbf{q}$ is known to be $(f | j_\nu | 0)$, then the matrix element of the electromagnetic interaction energy H' corresponding to the production of a free meson by an electron is

$$H' = A_\nu j_\nu. \quad (2)$$

The four-potential corresponding to the transition of the electron is a solution of the inhomogeneous wave equation

$$\square^2 A_\nu = -s_\nu,$$

where s_ν is the matrix element of the current density of this transition. It may readily be shown that the appropriate solution is

$$A_\nu = \frac{e\alpha_\nu}{k^2 - \epsilon^2} e^{i(\mathbf{k} \cdot \mathbf{x}) - i\epsilon t}. \quad (3)$$

In this equation $\mathbf{k}$ and ϵ are respectively the momentum and energy to be transferred from the electron to the system consisting of nucleons and mesons. If the nucleons are assumed to be of infinite mass, ϵ will then be the energy transferred to the free meson, i.e. $\epsilon^2 = \mu^2 + q^2$, where μ is the rest energy of the meson. The quantity α_ν occurring in the amplitude of this wave is the expectation value of the Dirac operator of the velocity $(\alpha_x, \alpha_y, \alpha_z, i)$. It will be seen from equation (2) that there is a strong tendency for the electromagnetic field to be propagated in the direction of the incident electron (the 'forward' direction) provided that the energy in both the final and initial state of the electron is large in comparison with the rest energy of the electron, which is unity in this system of units. The matrix elements j_ν and A_ν satisfy the equations

$$\frac{\partial j_\nu}{\partial x_\nu} = 0, \quad \frac{\partial A_\nu}{\partial x_\nu} = 0. \quad (4)$$

The first of these equations is the equation of continuity and the second the Lorentz condition. Making use of these relations we may transform equation (2) to the form

$$H' = \mathbf{A} \cdot \mathbf{j} - \frac{1}{\epsilon^2} (\mathbf{A} \cdot \mathbf{k})(\mathbf{k} \cdot \mathbf{j}) \quad (5)$$

by eliminating the fourth components of A_ν and j_ν .

Of the quantities involved in this equation, $\mathbf{A}$ is the only one which depends strongly on $\mathbf{k}$. As a function of $\mathbf{k}$, $\mathbf{A}$ shows a pronounced maximum when $\mathbf{k}$ is parallel to $\mathbf{p}$, in which case k is approximately equal to ϵ . Whenever $\mathbf{k}$ occurs in any quantity other than $\mathbf{A}$ we may therefore write $\mathbf{k} \simeq \epsilon \mathbf{p}/p$. With this approximation, involving an error of the order $1/p'$, the right-hand side of equation (5) reduces to the transverse part of the interaction between the vector potential and the transition current. We may therefore replace equation (5) by

$$H' = (\mathbf{A}_\perp \cdot \mathbf{j}_\perp), \quad (6)$$

where $\mathbf{F}_\perp$ denotes the transverse part of a vector field $\mathbf{F}$ defined by the relation

$$\mathbf{F}_\perp = \mathbf{F} - \frac{\mathbf{k}}{k^2}(\mathbf{k} \cdot \mathbf{F}).$$

The whole calculation then reduces to the determination of the matrix element $\mathbf{j}$ corresponding to the production of a free meson.

We calculate the matrix element $\mathbf{j}$ for the charged pseudoscalar meson theory, with the additional assumption that the nucleons are infinitely heavy (cf. Chang (1942) and others). We may therefore consider only the interaction of the meson field with the 'big' axial vector source function $\mathbf{P}$ of the nucleons. The Lagrangian function of the system of nucleons and mesons may then be written in the form

$$\mathfrak{L} = -\{(\nabla\varphi^\dagger \cdot \nabla\varphi) - \dot{\varphi}^\dagger\dot{\varphi} + \mu^2\varphi^\dagger\varphi + (\mathbf{P} \cdot \nabla\varphi^\dagger) + (\mathbf{P}^\dagger \cdot \nabla\varphi)\}. \quad (7)$$

Here φ and $\varphi^\dagger$ are the meson operators, $\varphi^\dagger$ corresponding to the production of a positive or the annihilation of a negative meson. To ensure gauge invariance it has to be assumed that under a gauge transformation $\mathbf{P}$ transforms in exactly the same way as φ . The only non-vanishing matrix elements of $\mathbf{P}$ then represent transitions from a proton into a neutron state. The expression (7) is the Lagrangian function in the absence of an external field; in the presence of an electromagnetic field we replace the operators $\partial/\partial t$ and ∇ by $\partial/\partial t - ieA_0$ and $\nabla - ie\mathbf{A}$ when acting on φ and by their complex conjugates when acting on $\varphi^\dagger$. With the usual procedure we obtain for the current vector

$$\mathbf{j} = \frac{\partial\mathfrak{L}}{\partial\mathbf{A}} = ie\{(\nabla\varphi^\dagger) \varphi - \varphi^\dagger(\nabla\varphi) + P^\dagger\varphi - P\varphi^\dagger\}. \quad (8)$$

This expression for the current has an obvious physical interpretation. In a simplified physical model in which a proton is considered as a mixture of a bare proton state and a state consisting of a neutron and a bound positive meson the production of a free meson may arise in either of two ways:

- (i) The meson is produced spontaneously by the source, i.e. the bare proton is transmuted to a bare neutron and a free positive meson; or
- (ii) the meson bound to the neutron is liberated by the action of the electromagnetic field.

Whereas the process (i) has no obvious counterpart in classical theory, (ii) bears a close resemblance to ionization.

For the nucleon-meson system the commutation relations for the quantities φ can be satisfied by putting

$$\varphi = \varphi_{\text{rad.}} = \sum_{\mathbf{q}} \sqrt{\left(\frac{1}{2e}\right)} \{a(\mathbf{q}) \phi_{\mathbf{q}}(\mathbf{r}) + b^\dagger(\mathbf{q}) \phi_{\mathbf{q}}^*(\mathbf{r})\}, \quad (9)$$

where $\phi_{\mathbf{q}}(\mathbf{r})$ are 'wave functions' with the asymptotic behaviour

$$\phi_{\mathbf{q}}(\mathbf{r}) \sim e^{i(\mathbf{q} \cdot \mathbf{r})}.$$

In the neighbourhood of the origin this simple form will be distorted as the result of the electrostatic interaction between the mesons and the nucleus. This gives rise to a 'Coulomb factor' which, in the first instance, we shall assume to be unity and

which we shall reintroduce at a later stage of the discussion. The operators a and b correspond to the annihilation of positive and negative mesons respectively, and for such transitions they have the value unity. Although they satisfy the commutation relations the expressions (9) do not, in the presence of sources, lead to a diagonal form for the Hamiltonian matrix. If we neglect the reaction of the meson field on the sources (i.e. if the source function P is assumed known) the general solution of the problem may be written in the form

$$\varphi = \varphi_0 + \varphi_{\text{rad.}}, \quad (10)$$

where φ_0 is a solution of the static equation

$$(\nabla^2 - \mu^2) \varphi_0 = -\nabla P. \quad (11)$$

It is then easily verified that equation (10) gives a function which satisfies the necessary commutation relations and makes the Hamiltonian diagonal.

To obtain a solution we must now specify the source function P of equation (11). In general we have

$$P = \frac{f}{\mu} \sigma \frac{\tau_1 - i\tau_2}{2}, \quad (12)$$

where σ is the spin operator acting on the nucleons and the τ 's are Pauli's isotopic spin matrices. The constant f is the coupling parameter which, in our system of units, is of the order unity. We now have to determine the matrix element corresponding to a transition in which one of the nuclear protons is transformed into a neutron and a momentum ξ is transferred to the nuclear system. We may then write

$$P = \frac{f}{\mu} e^{i\xi \cdot r} \sigma, \quad (13)$$

in which σ is the expectation value of the matrix element $\frac{1}{2}(\tau_1 - i\tau_2)$, that is,

$$\sigma = \sum_i \int dx_1 \dots dx_A u_f^*(x_1, \dots, x_A) \left(\sigma \frac{\tau_1 - i\tau_2}{2} \right)_i u_0(x_1, \dots, x_A)$$

In this integral the quantities x_i denote *all* the co-ordinates (including spin and isotopic spin) of the nuclear particles, and u_0 and u_f denote the anti-symmetrical wave functions of the initial and final states respectively. The solution of equation (11) is elementary

$$\varphi_0 = i(\xi \cdot \sigma) \frac{f}{\mu e'^2} e^{i\xi \cdot r}, \quad (14)$$

with $e'^2 = \mu^2 + \xi^2$. This method of building up solutions based on the static equation is justifiable only if the recoil energy of the nucleons can be treated as negligible.

Substituting from equations (9), (10) and (14) into the expression (8) for the matrix element of the current we obtain

$$\mathbf{j} = \frac{ie}{\sqrt{(2e)}\mu} \left\{ \frac{(\mathbf{q} + \xi)(\xi \cdot \sigma)}{\mu^2 + \xi^2} - \sigma \right\} e^{-i(\mathbf{q} - \mathbf{p}) \cdot \mathbf{r}}. \quad (15)$$

From equation (2) we have to combine this current with the matrix element of the vector potential corresponding to a transition in which the momentum k is trans-

ferred from the electron to the system of mesons and nucleons. From the conservation of momentum we obtain

$$\mathbf{q} - \boldsymbol{\xi} = \mathbf{k},$$

so that for the transverse component of the current we obtain

$$\mathbf{j}_\perp = \frac{ief}{\mu\sqrt{(2\epsilon)}} \left\{ \frac{2\mathbf{q}(\boldsymbol{\xi} \cdot \boldsymbol{\sigma})}{\mu^2 + \xi^2} - \boldsymbol{\sigma} \right\}. \quad (16)$$

Here again the first term is the 'ionization' term, and we see that it gives rise to an anisotropic distribution of mesons of very high energy with a marked preference for the forward direction. The second term due to the direct conversion of a quantum into a meson gives an isotropic distribution of mesons. When added together the two terms tend to cancel each other in the forward direction. It will be seen later that the major contribution at all meson energies is given by the second term in equation (16).

Inserting from equation (16) into (6) we obtain an expression for the matrix element H' which in Born's approximation is related to the differential cross-section, $\overline{d\varphi}$, by the formula

$$\overline{d\varphi} = \frac{\pi}{(2\pi)^6} \overline{|H'|^2} p'^2 q e d\Omega d\Omega'. \quad (17)$$

In this equation $\overline{|H'|^2}$ is the sum of $|H'|^2$ over the spins of the electron in the initial and final states and $1/(2J' + 1)$ times the sum over the degenerate initial and final states of the nucleus. J' is the quantum number of total spin for the final state of the nucleus. $d\Omega$ is the element of solid angle for the momentum vector $\mathbf{q}$ and $d\Omega'$ that for the electron.

This formula has been derived without account being taken of the effects of radiation damping. In the energy region in which the approximations we have already made hold good (i.e. up to energies of the order 600 MeV) these effects are, however, negligible (cf. Heitler 1949).

From equations (6), (16) and (17) we therefore obtain as a result of carrying out the summations over the nuclear states

$$\overline{d\varphi} = \frac{2\sigma_x^2}{(2\pi)^3 (2J' + 1) (137\mu)^2} f^2 deqp'^2 d\Omega d\Omega' (k^2 - \epsilon^2)^{-2} \left\{ \alpha_\perp \left(\frac{2q\xi}{\epsilon'^2} - I \right) \left(\frac{2q\xi}{\epsilon'^2} - I \right) \alpha_\perp \right\},$$

where the spin summation over the tensor $\alpha_\perp \alpha_\perp$ has yet to be carried out. Here σ_x denotes an arbitrary component of $\boldsymbol{\sigma}$. It will be observed from this expression that the 'ionization term' tends to cancel the leading term for mesons in the forward direction.

In the calculation of the total cross-section we first carry out the integration over the directions of the meson and obtain

$$\int_\Omega \overline{d\varphi} = \frac{\sigma_x^2}{\pi^2 (2J' + 1) (137\mu)^2} f^2 deqp'^2 d\Omega' (k^2 - \epsilon^2)^{-2} \left\{ 1 - \frac{2\mu^2}{eq} \left(\log \frac{\epsilon + q}{\mu} - \frac{q}{\epsilon} \right) \right\} (\alpha_\perp \cdot \alpha_\perp), \quad (18)$$

the dot denoting a scalar product. In performing this integration it was assumed that in the meson terms $k = \epsilon$, with the same degree of approximation as is used in the derivation of equation (6). The first term in the curly bracket is due to the

process (i), while the second comes from (ii) and the interference between the processes (i) and (ii). The second term is usually small; for non-relativistic mesons $q \ll \mu$, and this term behaves like q^2/μ^2 , while in the relativistic region it is of the order $2\mu^2/\epsilon^2$ times a slowly increasing logarithmic factor. Even in the case $\epsilon = 2\mu$ it amounts to only 26 % of the leading term. It can easily be shown (Sneddon & Touschek 1948) that

$$EE'(\alpha_{\perp}, \alpha_{\perp}) = -2 + \frac{E^2 + E'^2}{\epsilon^2} (k^2 - \epsilon^2) - \frac{1}{2\epsilon^2} (k^2 - \epsilon^2)^2, \quad (19)$$

where E and $E' = E - \epsilon$ are the energies of the electron in the initial and final states respectively. The integrations over the directions of the electron are now elementary and yield

$$\int_{\Omega} \int_{\Omega'} \overline{d\varphi} = \frac{4\sigma_x^2 f^2}{\pi(2J' + 1)(137\mu)^2} S(\epsilon) d\epsilon, \quad (20)$$

in which $S(\epsilon)d\epsilon$, the spectral distribution function of the mesons, is given by

$$S(\epsilon)d\epsilon = \frac{q d\epsilon}{\epsilon E} \left(\frac{E^2 + E'^2}{E^2} \right) \left\{ \log \frac{2EE'}{\epsilon} - \frac{2EE'}{E^2 + E'^2} \right\} \left\{ 1 - \frac{2\mu^2}{\epsilon q} \left(\log \frac{\epsilon + q}{\mu} - \frac{q}{\epsilon} \right) \right\}, \quad (21)$$

from which the total cross-section is obtained by integrating over ϵ .

3. NUMERICAL RESULTS AND DISCUSSION

In the derivation of equation (21) the effect of the electrostatic field of the nucleus has been neglected. We can correct this by multiplying $S(\epsilon)$ by N^2 —a 'Gamow factor' for the meson wave. This factor is defined by $N^2 = |\phi(0)|^2/|\phi(\infty)|^2$, where $\phi(r)$ is a regular solution of the meson wave equation in the electrostatic field of the nucleus. The nearly isotropic distribution of the mesons in the final state suggests that there is no appreciable error involved in taking s -waves for the meson functions. For light nuclei ($Z < 20$), for both positive and negative mesons N^2 can be taken to be unity. For large values of Z we can make use of the W.B.K. method to obtain an estimate of the function $\phi(r)$ and hence of N^2 . In the case of negative mesons it is easily shown that

$$N^2 = N_-^2 = q_s^{(-)}/q, \quad (22)$$

where $q_s^{(-)}$ denotes the momentum of the meson at the surface of the nucleus, so that

$$q_s^{(\pm)} = \{(\epsilon \mp e^2 Z/R)^2 - \mu^2\}^{\frac{1}{2}}.$$

When the meson has positive charge we must distinguish between two cases. In the first instance the energy of the meson may be less than the Coulomb barrier

$$\epsilon - \mu < \frac{e^2 Z}{R},$$

in which case N^2 is very small, so that in first approximation we may put $N^2 = 0$, since the short wave-length of the meson will make the potential barrier practically impenetrable. For higher energies ($\epsilon - \mu > e^2 Z/R$) we may take

$$N^2 = N_+^2 = q_s^{(+)} / q, \quad (23)$$

with g_s defined as above. The spectral distributions for positive and negative mesons $S_{\pm}(\epsilon) = N_{\pm}^2 S(\epsilon)$, obtained by multiplying (21) by the factors (22) and (23) respectively, are shown graphically in figure 1, for the case in which $e^2 Z/R \simeq 30$ (i.e. for a very heavy nucleus) and a maximum energy of 600. The meson mass is assumed to be 300 throughout. If we denote by n_+ and n_- respectively the total number of positive and negative mesons produced under these conditions we find by a numerical integration that

$$\frac{n_-}{n_+} = 1.74, \quad \frac{n_-}{n_0} = 1.27, \quad \frac{n_+}{n_0} = 0.73,$$

n_0 being proportional to the area under the dotted curve. The dotted curve gives the spectral distribution $S(\epsilon)$ defined by equation (21). Even in the case of this extremely heavy nucleus, the error involved in taking $N^2 = 1$ is less than 30 % for this energy. For higher energies the error is less; for lower energies it is greater.

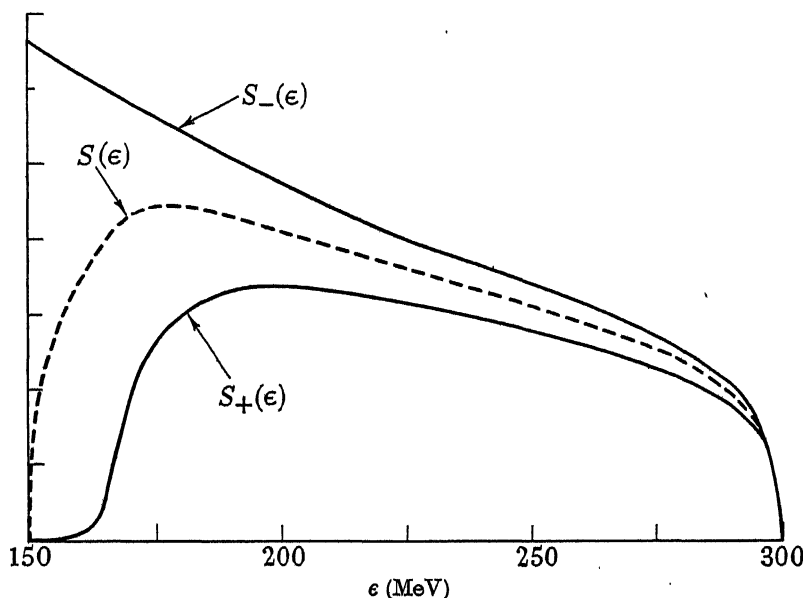


FIGURE 1. The distribution of energy of the mesons emitted from a very heavy nucleus when the energy of the incident electrons is 300 MeV. Arbitrary units are used for $S(\epsilon)$, the curve $S_-(\epsilon)$ applying when the liberated electrons have negative charge and the curve $S_+(\epsilon)$ when they have positive charge. The dotted curve $S(\epsilon)$ corresponds to the case $N^2 = 1$ (equation (21)), in which the electrostatic forces between meson and nucleus are neglected.

With this assumption we now proceed to estimate the total cross-section. For this total cross-section we find from equations (20) and (21)

$$\Phi = \Phi_0 g(E) \overline{\sigma_x^2 (2J' + 1)^{-1}}, \quad (24)$$

where

$$\Phi_0 = \frac{4f^2}{\pi(137\mu)^2} \quad (25)$$

and

$$g(E) = \int_{\mu}^E S(\epsilon) d\epsilon. \quad (26)$$

With $\mu = 300$ and $f \simeq 1$ we find that $\Phi_0 = 1.1 \times 10^{-30} \text{ cm}^2$. If the energy of both electron and meson in the final state is large compared to some average nuclear excitation energy (of the order of 10 MeV), we may consider the possible final states of the nucleus to be energetically equivalent. Then, making use of the completeness relation for the nuclear eigenfunctions, the sum over the possible final states of the nucleus becomes

$$\sum_f \frac{\sigma_x^2}{2J+1} = \begin{cases} Z, & \text{for positive mesons,} \\ A-Z, & \text{for negative mesons,} \end{cases} \quad (27)$$

The integral $g(E)$ can be estimated in the non-relativistic and extreme-relativistic approximations. From what has been said previously we may neglect the contribution to this integral made by the 'ionization term'. Replacing the logarithmic term in the integrand by a suitably chosen mean value ($E' = \frac{1}{2}(E - \mu)$) we find

$$g(E) \sim \frac{2\sqrt{2}}{3} \left(\frac{E - \mu}{\mu} \right)^{\frac{1}{2}} \log(E - \mu) \quad (\text{N.R.}) \quad (28)$$

$$\text{and} \quad g(E) \sim \frac{4}{3} \log(2E) \quad (\text{E.R.}). \quad (29)$$

As a result of a numerical integration of the exact spectral distribution function (21) we find that $g(600) = 1.9$. This is a factor four less than the value given by either (28) or (29), but we would not expect agreement in this case, since the energy considered is outside the range of both approximations. In deriving these approximations we have neglected terms which would tend to decrease the total cross-section, so that (28) and (29) should be considered as defining upper limits.

It follows from these considerations that the cross-section is proportional to the number of nucleons in the target nucleus, a result which is immediately evident by considering the small wave-lengths of all particles occurring in the process. The absolute value of the cross-section is very small; at 300 MeV it is approximately $2 \times 10^{-30} \text{ cm}^2$ per nucleon. From (29) it is seen that the cross-section increases slowly with energy, but the result is unreliable in the high-energy region because we have neglected to take into account the effects of nuclear recoil and radiation damping. However, it follows from recent calculations reviewed by Heitler (1949) that up to energies of 600 MeV the influence of radiation damping is not large in the case of pseudoscalar mesons.

The cross-section for this process is only about one-hundredth part of the cross-section for the production of mesons by γ -rays, as calculated by Chang (without radiation damping, 1942) and Peng (Morette & Peng 1947). It should be observed, however, that the values given by Peng are rather high owing to the small meson mass assumed in the calculation. In order to observe the production of mesons by electrons it appears therefore that very thin targets (of the order of one hundredth of a radiation length) must be used. In the case of lead this is of the order of 0.05 mm.

The shape of the meson spectrum produced by electrons is roughly the same as that produced by a thin target 'Brems-spectrum' of γ -rays. The reason for this is that the nucleus radiates its mesons under the influence of the Maxwell field of an electron transition. The spectrum of this equivalent Maxwell field is approximately

the same as the spectrum derived from thin-target radiation. Even quantitatively we should expect the resemblance to be close, considering the equivalence of electrons and photons revealed in the Weizsäcker-Williams method.

REFERENCES

- Chang, T. S. 1942 *K. danske vidensk. Selsk. Mat.-fys. Medd.* **19**, 10.
 Heitler, W. 1949 *Rev. Mod. Phys.* **21**, 113.
 Møller, C. & Rosenfeld, L. 1940 *K. danske vidensk. Selsk. Mat.-fys. Medd.* **17**, 8.
 Morette, C. & Peng, H. W. 1947 *Nature*, **160**, 59.
 Sneddon, I. N. & Touschek, B. F. 1948 *Proc. Roy. Soc. A*, **193**, 344.
 Weizsäcker, C. F. v. 1934 *Z. Phys.* **88**, 612.
 Williams, E. J. 1934 *Phys. Rev.* **45**, 729.

Statistical mechanics and the partition of numbers

I. The transition of liquid helium

By H. N. V. TEMPERLEY, *King's College, University of Cambridge*

(Communicated by D. R. Hartree, F.R.S.—Received 18 February 1949—
 Revised 7 May 1949)

The existing theory of 'Bose-Einstein condensation' is compared with some results obtained from the theory of partition of numbers. Two models are examined, one in which the energy levels are all equally spaced, the other being the perfect gas model. It is concluded that orthodox theory can be relied upon at very high and at very low temperatures, also that the condensation phenomenon is a real one, but that it is not correctly described by orthodox theory, the position of the transition temperature and the form of the specific heat anomaly both being given wrongly.

1. INTRODUCTION

It has long been recognized that the central problem of statistical mechanics, the determination of the number of ways in which a given amount of energy can be shared out among the different possible states of an assembly, is essentially a problem of the same type as that of the determination of the number of partitions of a number into integral parts under certain restrictions. The application by Fowler and Darwin, in the former writer's great book *Statistical mechanics*, of the powerful methods of complex variable theory that had already proved so successful in analytic number theory, served to underline this affinity, which has been explicitly mentioned by a number of writers, for example, Auluck & Kothari (1946). In the present paper, some of the methods of partition theory will be applied to a study of the 'Bose-Einstein condensation' of a perfect gas. This problem has been chosen for study because a recent paper by Schubert (1946) has thrown doubt on the validity of orthodox theory. The difficulty arises when one attempts to use the Fowler-Darwin,

method of steepest descents to determine the coefficient of x^N in the generating function $\prod_r (1 - xz^{E_r})^{-1}$, where E_r is the energy of the r th state. In the Fowler-Darwin method, this is done by means of a contour integration with respect to the variable x , the contour being a circle of radius λ encircling the origin, λ being chosen so that the circle passes through the saddle-point of the integrand on the real axis. The results of orthodox theory now follow if it can be shown that the main contribution to the contour integral comes from a small region of the circle in the neighbourhood of the real axis. If so, it then follows that the occupation number of the r th state, of weight w_r , is given by the well-known expression $\bar{N}_r = \frac{w_r \lambda e^{-E_r/kT}}{1 - \lambda e^{-E_r/kT}}$, and, if this may be assumed, the detailed theory of the Bose-Einstein condensation can be rigorously proved as was done by London (1938), and Fowler & Jones (1938).

The contour used by Fowler and Darwin is the circle $x = \lambda e^{i\alpha}$, and their procedure is to expand the integrand in the form $\exp(\sum_r A_r \alpha^r)$, where A_1 vanishes because $x = \lambda$ is a saddle-point of the integrand. In general A_2 is large and negative, and it is possible to find a small value α_0 of α such that $A_2 \alpha_0^2$ has become numerically very large, while the terms involving $A_3, A_4, \dots$ are still negligibly small, on account of the fact that they involve higher powers of α . The results of orthodox theory now follow if it can be shown that:

(a) The remainder of the circle $\pi \geq |\alpha| \geq \alpha_0$ contributes to the integral an amount that is negligibly small compared with the contribution of the region of the circle $|\alpha| \leq \alpha_0$.

(b) The expansion of the integrand as the exponential of a power series in α is convergent in the region $|\alpha| \leq \alpha_0$.

Assumption (a) does not seem to have been critically examined, but Schrödinger (*Statistical thermodynamics*) has examined at length the analogous case of integration over the z -variable in classical statistics, and has proved that here the analogous assumption introduces no physically significant error. Assumption (b) has been examined by Schubert (1946) and by Fraser (unpublished) for the perfect gas model, who find that, for temperatures in the region of the alleged transition temperature, the expansion of the integrand used by Fowler and Darwin fails to converge in part of the range of $|\alpha| \leq \alpha_0$. The reason for this is that the integrand has a singularity at $\lambda = 1$, so that when the saddle-point approaches the unit circle, as it does in this temperature region, the circle of convergence of the expansion of the integrand becomes insufficient to include the whole of the range of $|\alpha| \leq \alpha_0$, and the Fowler-Darwin method breaks down. It is, however, still possible that the orthodox formula

$$\bar{N}_r = \frac{w_r \lambda e^{-E_r/kT}}{1 - \lambda e^{-E_r/kT}}$$

is universally true, and it is the purpose of this paper to examine this question by methods which do not involve contour integration over the variable x .

A recent attempt by Leibfried (1947) to show that the orthodox formula for $\bar{N}_r$ is always valid for the perfect gas model must be rejected as completely fallacious, the fallacy being that one cannot differentiate an asymptotic expansion without a careful examination of the remainder term, still less may one differentiate the

leading term after rejecting all the others, as Leibfried attempts to do. Because each term is negligible compared with the one before it, it does not follow that the same is true of their derivatives.

The investigation now to be described seems to show conclusively that the orthodox expression for the occupation number $\bar{N}_r$ is not universally valid and that the true distribution law is more complicated. The orthodox law appears to be valid at very high and at very low temperatures, but for the two distributions of energy level examined there is a region of intermediate temperatures where it fails. It is not yet possible to say whether this failure always occurs in Bose-Einstein statistics, or whether it is confined to models which would, on orthodox theory, show a condensation phenomenon.

The two models studied are, first, one in which the energy-levels are equally spaced. The consequences of this model are examined, first, if the orthodox expression for the occupation number is assumed, secondly, making no such assumption but using the analytic theory of the partition of numbers. It is concluded that the orthodox expression gives incorrect results in the transition region (in particular that there is no 'sharp' specific heat anomaly). The second model studied is the more complicated one of the perfect gas, which cannot yet be solved so completely by partition methods, but again the orthodox expression is shown to be invalid for a certain range of temperatures.

2. A MODEL WITH EQUALLY SPACED LEVELS

This model has already been considered by Auluck & Kothari (1946) in connexion with their work on the partition of numbers, the problem they consider being the 'radiation problem', that is, the sharing of a fixed amount of energy between N harmonic oscillators of equal frequency. We shall consider the rather different problem of N particles obeying Bose-Einstein statistics distributed among infinitely many energy-levels $0, \epsilon, 2\epsilon, 3\epsilon, \dots$ of uniform spacing ϵ , in such a way that the total energy is ϵE . Auluck & Kothari find expressions for $p_N(E)$, the number of ways of dividing an integer E into N or less integral parts, from the statistical mechanics of harmonic oscillators, and obtain results agreeing with those obtained for $p_N(E)$ by purely mathematical methods, so that for the radiation problem the use of the method of steepest descents for the evaluation of their integrals seems justified. In our problem, we shall obtain results differing widely from those given by orthodox theory.

The generating function for $p_N(E)$ is

$$\sum_N \sum_E p_N(E) x^N z^E = 1/(1-x)(1-xz)(1-xz^2)(1-xz^3)\dots \quad (1)$$

In our problem, if $\bar{N}_r$ is the mean occupation number of the state with energy $r\epsilon$ referred to the lowest state as zero, the generating function for $\bar{N}_r p_N(E)$ is easily seen to be

$$\frac{xz^r}{1-xz^r(1-x)(1-xz)(1-xz^2)\dots} \quad (2)$$

According to orthodox theory we now apply the method of contour integration with steepest descents to pick out the appropriate powers of x and z in these two expressions, and obtain for the ratio the expression

$$\bar{N}_r = \frac{\lambda \theta^{re}}{1 - \lambda \theta^{re}}, \quad (3)$$

where λ and θ^e are the values of x and z respectively at the saddle-point.

In this particular model, the coefficient of x^N can be picked out exactly, so that the contour integration with respect to x is not really necessary. We shall first assume the accuracy of equation (3), and that the spacing of the levels ϵ is independent of N . In these circumstances we obtain a discontinuity in the specific heat at the transition temperature. (If we wanted to discuss the 'two-dimensional Bose-Einstein gas', which can be approximated to by a spacing of energy-levels at equal intervals, we should have to allow ϵ to tend to zero like $1/A$, where A is the area within which the molecules are confined. For a constant surface density of particles N/A , this would make ϵ tend to zero like $1/N$ as the assembly became very large, and, in these circumstances it is well known that no discontinuity appears.)

We shall then compare these results with those obtained by a method that does not assume the validity of equation (3), and does not necessitate the use of a contour integral to determine the coefficient of x^N . In Schubert's (1946) discussion of a perfect gas, it was the corresponding contour integration that was found to be impossible by the method of steepest descents, and we shall find evidence of a similar failure here, which may be traced to the fact that, if N is large compared with $E^\dagger$, the function $p_N(E)$ varies so slowly with N that it is not surprising that the method of steepest descents fails. On the other hand, $p_N(E)$ always increases very rapidly with E , and the contour integration with respect to z , to determine the coefficient of z^E , seems to be always possible by the method of steepest descents. A corresponding result is implicitly assumed by Schubert (1946) for the perfect gas.

The result of the amended treatment of the equally spaced model is that there is certainly a range of temperatures just above absolute zero for which the occupation numbers of all states but the lowest are independent of N , and that there is a transition temperature above which the fact that N is finite begins to affect the occupation numbers, but that this effect certainly cannot be described by a *single* parameter λ in the manner of equation (3). Furthermore, this transition temperature is much lower than that predicted by orthodox theory based on equation (3). It seems probable that the transition is of an unfamiliar type, in that it is not associated with a discontinuity of any finite derivative of the partition function, but our present knowledge of the behaviour of the function $p_N(E)$ is not sufficient to enable us to be quite certain of this.

Finally, some results are obtained for the perfect gas model. It is first shown that, in the 'classical' limiting case of high temperatures, the consequences of equation (3) are practically equivalent to some well-known results on the representation of large numbers as sums of integral squares. Secondly, it is shown that equation (3) is correct at sufficiently low temperatures, but that the transition temperature again lies much lower than it would according to orthodox theory, so that the transition temperature

for a perfect gas of particles of mass equal to that of a helium atom would lie well below 1°K . If, however, the assumption of a perfect gas were replaced by the more realistic assumption that a typical helium atom is confined within a 'cage' of dimensions of the order of magnitude of the atomic distances in liquid helium, then *this* model would be expected to have a transition temperature of the order of 1 to 2°K .

3. A MODEL WITH EQUALLY SPACED ENERGY-LEVELS. ORTHODOX TREATMENT

We consider an assembly containing infinitely many energy-levels spaced at equal intervals ϵ above the lowest one, each being capable of being occupied any number of times. We let the total number of particles be N and the total energy ϵE . It is not claimed that this model represents any actual physical situation, but it does exhibit in what way the orthodox and exact treatments fail to agree. According to the orthodox theory we have

$$N = \sum_{r=0}^{\infty} \frac{\lambda \theta^{r\epsilon}}{1 - \lambda \theta^{r\epsilon}}, \quad E = \sum_{r=0}^{\infty} \frac{r \lambda \theta^{r\epsilon}}{1 - \lambda \theta^{r\epsilon}}. \quad (4)$$

If we keep ϵ finite, it is never possible to represent these sums as tending to integrals, but they can be summed without much difficulty by Poisson's formula, which enables one to sum a series if one can sum the series formed by applying the Fourier transform to each term. We apply this process to the series for N after removing the first term, and to the series for E . We obtain, putting $\lambda = 1/A$, $\theta = e^{-1/kT}$,

$$\begin{aligned} N - \frac{1}{A-1} - \frac{1}{2} \frac{1}{A e^{\epsilon/kT} - 1} &= \sum_{l=-\infty}^{+\infty} \int_0^{\infty} \frac{e^{2\pi i l t}}{A e^{\epsilon(1+l)/kT} - 1} dt \\ &= \sum_{s=1}^{\infty} \sum_{l=-\infty}^{+\infty} \int_0^{\infty} e^{2\pi i l t} A^{-s} e^{-s(t+1)\epsilon/kT} dt \\ &= \sum_{s=1}^{\infty} \sum_{l=-\infty}^{+\infty} \frac{A^{-s} e^{-s\epsilon/kT}}{(s\epsilon/kT) + 2\pi i l} \\ &= \frac{kT}{\epsilon} \sum_{s=1}^{\infty} \frac{A^{-s} e^{-s\epsilon/kT}}{s} + \frac{2\epsilon}{kT} \sum_{s=1}^{\infty} \sum_{l=1}^{\infty} \frac{s A^{-s} e^{-s\epsilon/kT}}{(s\epsilon/kT)^2 + 4\pi^2 l^2} \\ &= -\frac{kT}{\epsilon} \log \left(1 - \frac{e^{-\epsilon/kT}}{A} \right) + \frac{\epsilon}{kT} \frac{1/A e^{\epsilon/kT}}{(1 - 1/A e^{\epsilon/kT})^2} O(1). \end{aligned} \quad (5)$$

The corresponding equation for E is

$$\begin{aligned} E - \frac{1}{2} \frac{1}{A e^{\epsilon/kT} - 1} &= \sum_{l=-\infty}^{+\infty} \int_0^{\infty} \frac{(t+1) e^{2\pi i l t}}{A e^{\epsilon(t+1)/kT} - 1} dt \\ &= \sum_{s=1}^{\infty} \sum_{l=-\infty}^{+\infty} \int_0^{\infty} (t+1) e^{2\pi i l t} A^{-s} e^{-s(t+1)\epsilon/kT} dt \\ &= \sum_{s=1}^{\infty} \sum_{l=-\infty}^{+\infty} \frac{A^{-s} e^{-s\epsilon/kT}}{(s\epsilon/kT) + 2\pi i l} + \sum_{s=1}^{\infty} \sum_{l=-\infty}^{+\infty} \frac{A^{-s} e^{-s\epsilon/kT}}{[(s\epsilon/kT) + 2\pi i l]^2} \\ &= \frac{kT}{\epsilon} \sum_{s=1}^{\infty} \frac{A^{-s} e^{-s\epsilon/kT}}{s} + \left(\frac{kT}{\epsilon} \right)^2 \sum_{s=1}^{\infty} \frac{A^{-s} e^{-s\epsilon/kT}}{s^2} \\ &\quad + \frac{2\epsilon}{kT} \sum_{s=1}^{\infty} \sum_{l=1}^{\infty} \frac{s A^{-s} e^{-s\epsilon/kT}}{(s\epsilon/kT)^2 + 4\pi^2 l^2} + 2 \sum_{s=1}^{\infty} \sum_{l=1}^{\infty} \frac{A^{-s} e^{-s\epsilon/kT} [(s\epsilon/kT)^2 - 4\pi^2 l^2]}{[(s\epsilon/kT)^2 + 4\pi^2 l^2]^2}. \end{aligned} \quad (6)$$

We know that A may never be less than unity because this would make the occupation number of the lowest state infinite. Equation (6) shows that, to make E a large number, we must have $\epsilon/kT = O(1)$. The third and fourth terms in equation (6) (right-hand side) are then at most $O(kT/\epsilon)$, as is the term $\frac{1}{2}(A e^{\epsilon/kT} - 1)^{-1}$, while the first term on the right-hand side is $O\left\{\frac{kT}{\epsilon} \log\left(\frac{kT}{\epsilon}\right)\right\}$ and the second term is $O\left(\frac{kT}{\epsilon}\right)^2$. On the right-hand side of equation (5), the second term can only become large if A is nearly unity, in which case it is of order kT/ϵ , but the first term is then of order $\frac{kT}{\epsilon} \log\left(\frac{kT}{\epsilon}\right)$. If, for example, we have A equal to $1 + \frac{O(1)}{N^{\frac{1}{2}}}$, then, in the limiting case of a very large assembly, the second and third terms on the left-hand side are negligible compared with N , and we shall have

$$N \sim \frac{kT}{\epsilon} \log\left(\frac{kT}{\epsilon}\right), \quad E \sim \left(\frac{kT}{\epsilon}\right)^2 \sum_{s=1}^{\infty} \frac{1}{s^2} = \frac{\pi^2}{6} \left(\frac{kT}{\epsilon}\right)^2. \quad (7)$$

If now A takes the slightly smaller value $1 + \frac{O(1)}{N}$, the expression for E is practically unaffected, but the term $1/(A-1)$ in equation (5) now contributes effectively to N . Thus, we have a situation very similar to that in the orthodox theory of the perfect gas, the energy at a given temperature, being no longer affected by N when this number rises above the value $\frac{\sqrt{6}}{\pi} E^{\frac{1}{2}} \log E$, any further increase merely affecting A , but, as the change in A is only of the order of $1/N$, the only term in equations (6) and (7) to be significantly affected is the term $1/(A-1)$, which determines the occupation number of the very lowest level.

The nature of the anomaly that occurs can be found by differentiating the expression (4) for E with respect to temperature, and summing the resulting series for $\partial E/\partial T$, by means of Poisson's formula. This gives an expression for the specific heat in terms of $\partial A/\partial T$, and the latter quantity can be evaluated by differentiating the series (4) for N with respect to temperature, using the fact that N is constant, so that the series for $\partial N/\partial T$ must vanish. Both these series can be summed by the same process as that used for equations (5) and (6), that of picking out the dominant terms in the transformed series. Below the transition temperature the quantity $\partial A/\partial T$ is negligibly small, while just above the transition temperature it is, in the limiting case of very large N and E , given by the simple expression $1/T_0$, where T_0 is the transition temperature as determined by equation (7). The discontinuity in $\partial A/\partial T$ leads to a negative discontinuity in the specific heat, which drops discontinuously from the value $\frac{\pi^2 k^2}{3 \epsilon}$ just below the transition temperature to the value $\left(\frac{\pi^2}{3} - 1\right) \frac{k^2}{\epsilon}$ just above this temperature. The analysis is rather lengthy on account of the fact that both the transformed series contain a great many terms whose orders of magnitude all have to be examined, so it is omitted, as it contains no features of any special interest.

This model thus gives a discontinuity in the specific heat itself, in contrast with the discontinuity in the derivative of the specific heat found for the perfect gas. The

slightly more drastic nature of the discontinuity is due to the fact that we are supposing the separation of the energy-levels to remain finite, while in the perfect gas, the spacing of the levels becomes zero for a very large assembly.

4. A MORE EXACT TREATMENT OF THE MODEL WITH EQUALLY SPACED ENERGY-LEVELS

In this treatment we abandon equation (3) and return to the exact equations (1) and (2). For this simple model, one of the integrations by the method of steepest descents can be avoided by algebraic means. In equation (1), putting xz for x we obtain

$$\sum_N \sum_E p_N(E) x^N z^{E+N} = 1/(1-xz)(1-xz^2)(1-xz^3)\dots = (1-x) \sum_N \sum_E p_N(E) x^N z^E. \quad (8)$$

Equating coefficients of x^N , we find

$$\sum_E p_{N-1}(E) z^E = (1-z^N) \sum_E p_N(E) z^E,$$

so that

$$\sum_E p_N(E) z^E = 1/(1-z)(1-z^2)\dots(1-z^N). \quad (9)$$

Expanding the first factor in expression (2) in powers of x , and remembering that expression (2) may also be written as $\sum_N \sum_E \bar{N}_r p_N(E) x^N z^E$, we obtain

$$\sum_E \bar{N}_r p_N(E) z^E = \frac{z^r(1-z^N) + z^{2r}(1-z^N)(1-z^{N-1}) + \dots}{(1-z)(1-z^2)\dots(1-z^N)}. \quad (10)$$

One can prove formally the applicability of the saddle-point method to determine the ratio of the coefficients of z^E in (9) and (10). An outline of this proof is as follows:

The contour integral $\oint \frac{dz}{z^E(1-z)(1-z^2)\dots}$ has been intensively studied, e.g. by Hardy's *Ramanujan* (1940, chap. VIII), using the following method of evaluating such integrals. A given radius R (< 1) of the contour is chosen, and the contour is divided into arcs each of which arcs is associated with one of the rational fractions in the Farey series of order n , this Farey series consisting of those rational fractions whose denominators are not greater than n . Each rational fraction p is associated with a point on the unit circle whose polar angle is $2\pi p$, which point is a singularity of $f(z)$. The function $f(z) = \frac{1}{(1-z)(1-z^2)\dots}$ is transformed by the formulae of elliptic

function theory, and it is shown that, in each arc of the contour, it is sufficiently accurate to replace $f(z)$ by its asymptotic form in the neighbourhood of the point on the unit circle associated with this arc, a different transformation formula holding for each arc, according to the particular polar angle $2\pi p$. The investigation shows that the main contribution to the contour integral is always given by the arc in the neighbourhood of the real axis associated with the singularity $z = 1$ of $f(z)$, in other words that, to use Hardy's language, this singularity of $f(z)$ is by far the 'heaviest' one. The contributions to the contour integral associated with the other rational singularities of $f(z)$ are together negligible compared with this one (however large

n may be). For any given value of n , it can further be shown that the error in the evaluation of the contour integral can even be made less than unity if n is made sufficiently large and all the corresponding singularities are considered.

Thus, for the contour integral containing $f(z)$, we can show quite rigorously, first that the only contribution of physical importance for a very large assembly comes from the immediate neighbourhood of the real axis, secondly that $f(z)$ can always be expressed in a form that converges on the whole of the arc corresponding to $z = 1$, so that both our conditions (a) and (b) above (p. 362) would be satisfied if the integrand contained $f(z)$. This situation is not changed by replacing $f(z)$ by the function

$\frac{1}{(1-z)(1-z^2)\dots(1-z^N)}$ which only means that the essential singularities of $f(z)$ on the unit circle are replaced by poles of order N , $\frac{1}{2}N$, $\frac{1}{3}N$, etc., nor does multiplication of the integrand by the function $z^r(1-z^N) + z^r(1-z^N)(1-z^{N-1}) + \dots$ (which has at most a simple pole at $z = 1$) affect the position of the saddle-point, because the integrand itself has a pole of order N at $z = 1$ and is therefore varying with extreme rapidity in this neighbourhood.

We thus obtain the following expression for $\bar{N}_r$,

$$\bar{N}_r = \theta^{er}(1 - \theta^{Ne}) + \theta^{2er}(1 - \theta^{Ne})(1 - \theta^{(N-1)e}) + \dots \quad (r \geq 1), \quad (11)$$

where θ^e is the value of z at the saddle-point. This argument fails for the very lowest state, $r = 0$, but, as a matter of fact, expression (11) is valid in this case also. To prove this, we have

$$\bar{N}_0 = N - \sum_{r=1}^{\infty} \bar{N}_r = N - \frac{\theta^e(1 - \theta^{Ne})}{1 - \theta^e} - \frac{\theta^{2e}(1 - \theta^{Ne})(1 - \theta^{(N-1)e})}{1 - \theta^{2e}} - \dots$$

(using equation (11) for $\bar{N}_r$ and summing from $r = 1$ to infinity), which gives

$$\bar{N}_0 = (1 - \theta^{Ne}) + (1 - \theta^{Ne})(1 - \theta^{(N-1)e}) + \dots$$

by setting $p = \theta^e$ in the identity

$$N \equiv \frac{1 - p^N}{1 - p} + \frac{(1 - p^N)(1 - p^{N-1})}{1 - p^2} + \dots$$

This identity may be proved as follows: Let k_N stand for the expression

$$k_N = \frac{1 - p^N}{1 - p} + \frac{(1 - p^N)(1 - p^{N-1})}{1 - p^2} + \dots$$

Then

$$\begin{aligned} k_N - k_{N-1} &= p^{N-1} + p^{N-2}(1 - p^{N-1}) + p^{N-3}(1 - p^{N-1})(1 - p^{N-2}) + \dots \\ &\quad + (1 - p^{N-1})(1 - p^{N-2}) \dots (1 - p) \\ &= p^{N-1} + (1 - p^{N-1})(k_{N-1} - k_{N-2}). \end{aligned}$$

Now $k_2 - k_1$ is identically unity, therefore so are $k_3 - k_2$, $k_4 - k_3$, etc., and since k_1 is unity it follows that k_N is identically equal to N and therefore that equation (11) is also true for $r = 0$.

For $N > E$, the terms in expression (10) that involve z^N and higher powers of z can contribute nothing to the coefficient of z^E , while, until N drops to a value com-

parable with $E^{\frac{1}{2}}$, they give a contribution to expression (11) that vanishes in comparison with the main terms. For N less than E , the expression for $\bar{N}_r$ therefore takes the form

$$\bar{N}_r \approx \theta^{er} + \theta^{2er} + \theta^{3er} + \dots = \frac{\theta^{er}}{1 - \theta^{er}}, \quad (12)$$

which agrees with the very low temperature expression for $\bar{N}_r$ according to orthodox theory provided that we make the customary identification $\theta = e^{-1/kT}$. For T very large, expression (11) agrees with the 'classical' expression $\theta^{er}/(1 + \theta^e + \theta^{2e} + \dots)$ for $\bar{N}_r$, provided that we can neglect θ^e in comparison with unity. The customary relation between θ and T therefore gives us results that agree with orthodox theory both for very high temperatures and for a finite range of energies above absolute zero, thus justifying this identification of θ .

We now examine whether expression (11) can be made to agree with orthodox theory in the region of the transition temperature predicted by the orthodox theory. We notice first of all the following relation between the values of $\bar{N}_r$ for two consecutive values of N , which follows easily from expression (11),

$$\bar{N}_r(N) = \theta^{er}(1 - \theta^{Ne})(1 + \bar{N}_r(N-1)). \quad (13)$$

If orthodox theory were correct in the transition region, then it would be possible to express the occupation numbers in the form

$$\bar{N}_r(N) = \frac{\lambda_{N,r} \theta^{re}}{1 - \lambda_{N,r} \theta^{re}}, \quad (14)$$

where $\lambda_{N,r}$ may be written $1 - \delta_{N,r}$, where $\delta_{N,r}$ is a small quantity which may depend on N , but whose leading term must be independent of r . Substituting expression (14) into equation (13), expressing the λ 's in terms of the δ 's, and neglecting products of the δ 's,

$$\delta_{N,r} = \theta^{Ne}(1 - \theta^{re}) + \theta^{re} \delta_{N-1,r} + O(\theta^{2Ne}), \quad (15)$$

which implies that the leading terms in $\delta_{N,r}$ must take the following form:

$$\begin{aligned} \delta_{N,r} &\approx \frac{(1 - \theta^{re})(\theta^{Ne} - \theta^{Nre})}{1 - \theta^{(r-1)e}} \quad (r \geq 2) \\ &\approx N\theta^{Ne}(1 - \theta^e) \quad (r = 1). \end{aligned} \quad (16)$$

Equation (16) shows clearly that $\delta_{N,r}$ is *not* independent of r , a different degeneracy parameter $\lambda_{N,r}$ appearing for each energy level, and it must be concluded that orthodox theory is not correct, at least not in the supposed transition region where $\lambda_{N,r}$ should be nearly unity.

Orthodox theory therefore cannot be used in the transition region, where we can point to two other discrepancies with the treatment we have just given. In the first place, the onset of the transition is not located correctly. Orthodox theory (equation (7)) predicts that the transition should take place when N is proportional to $E^{\frac{1}{2}} \log E$, while the exact theory shows that the transition must set in, in the sense of the occupation numbers of the state other than the lowest becoming functions of N as well as of E , when N becomes less than E . In the second place, orthodox theory pre-

dicts a jump in the specific heat, while the exact theory certainly predicts an anomaly, but not one corresponding to a phase-transition of any finite order. The following is the asymptotic expression for $p_N(E)$, which can be obtained from equation (9) by using an asymptotic expression for the function $1/(1-z)(1-z^2)\dots(1-z^N)$ in the neighbourhood of the singularity $z = 1$, as in the classical work on the partition of numbers of Hardy & Ramanujan (1918). It can also be obtained, as was done by Auluck & Kothari (1946), from the accepted statistical mechanical formulae for the distribution of energy among simple harmonic oscillators. Auluck, Chowla & Gupta (1942) have obtained a similar result by methods which do not use a contour integral at all.

$$p_N(E) \sim \frac{1}{4\sqrt{3}E} \exp \left[\pi \sqrt{\frac{2E}{3}} - \frac{\sqrt{(6E)}}{\pi} \exp \left(-\frac{\pi N}{\sqrt{(6E)}} \right) \right] \quad (N \gg E^{\frac{1}{2}}). \quad (17)$$

By the Boltzmann formula the entropy S is given by the expression $S = k \log p_N(E)$, which reduces to $k \log p(E)$ when $N \gg E$. The quantity $k \log \left(\frac{p_N(E)}{p(E)} \right)$ is therefore the defect in entropy caused by the fact that N is finite, which will lead to a specific heat anomaly as E passes through the value N . (The temperature is determined by the usual relation $T = \partial S / \partial E$.) The form of equation (17) suggests that none of the derivatives of S with respect to E or T are discontinuous as E passes through the value N , so that no discontinuity of co-operative type is to be expected. This result is not yet formally proved, on account of the difficulties of differentiating asymptotic expressions, but an elementary argument can be used to establish it. The numbers $p_E(E)$ and $p_{E-1}(E)$ only differ by unity, namely by the one partition in which E is divided into E parts, all of them unity. There is therefore, for E large, no significant discontinuity in the first derivative of $\log p_N(E)$. A similar argument can be applied to the second, third and higher derivatives by counting up the number of partitions that we lose in going from $p_{E-1}(E)$ to $p_{E-2}(E)$, $p_{E-3}(E)$, etc. It is therefore extremely likely that there is no actual discontinuity in any finite derivative of S . This situation seems to be rather like what one would observe supposing that one had a transition temperature T_c , with no anomaly in the entropy below T_c , but an extra term proportional to $\exp \left(\frac{B}{T_c - T} \right)$, where B is a constant, above T_c . Such an extra term would not give any discontinuity in the specific heat or in any of its derivatives, and, from an experimental point of view, would be indistinguishable from a 'smooth' specific heat anomaly.

Equation (17) is of interest from another point of view, in indicating a possible reason for the failure of the saddle-point method when applied to the variable x in equations (1) and (2), namely, that in the transition region the value of $p_N(E)$ is changing extremely slowly with N . This, combined with the fact that the saddle-point value of x is nearly unity means that, for a fixed E , expression (1) contains a very large number of terms $\sum_N p_N(E) x^N$ all of which are of the same order of magnitude, so that one cannot necessarily replace the logarithm of this series by the logarithm of a single term with sufficient approximation. Since $p_N(E)$ always varies extremely rapidly with the variable E , no such effect occurs to spoil the validity of the z -integration.

We have still to examine the behaviour of $\bar{N}_0$ according to the exact theory, as equation (11) gives no direct information on this point. Equation (10), however, is valid, giving us

$$\sum_E N_0 p_N(E) z^E = \frac{(1-z^N) + (1-z^N)(1-z^{N-1}) + \dots}{(1-z)(1-z^2)(1-z^3) \dots (1-z^N)}. \quad (18)$$

Let f_N stand for the series

$$(1-z^N) + (1-z^N)(1-z^{N-1}) + \dots$$

Then we have the following relation

$$f_N = (1-z^N)(1+f_{N-1}). \quad (19)$$

Now set $f_N = N - g_N$ and equation (19) becomes

$$g_N = Nz^N + (1-z^N)g_{N-1}. \quad (20)$$

If we now set

$$g_N = h_N(1-z)(1-z^2) \dots (1-z^N)$$

we obtain from equation (20)

$$h_N - h_{N-1} = \frac{Nz^N}{(1-z)(1-z^2) \dots (1-z^N)}, \quad (21)$$

which gives us the following expression for h_N :

$$h_N = \frac{Nz^N}{(1-z)(1-z^2) \dots (1-z^N)} + \frac{(N-1)z^{N-1}}{(1-z) \dots (1-z^{N-1})} + \dots + \frac{z}{1-z}. \quad (22)$$

Equation (18) may now be written

$$\sum_E \bar{N}_0 p_N(E) z^E = \frac{N}{(1-z)(1-z^2) \dots (1-z^N)} - h_N \quad \text{or} \quad \sum_E (N - \bar{N}_0) p_N(E) z^E = h_N, \quad (23)$$

since, from equation (9), $p_N(E)$ is the coefficient of z^E in $1/(1-z)(1-z^2) \dots (1-z^N)$. We already know that $p_N(E)$ becomes independent of N if $N \geq E$, and equation (22) shows that the coefficient of z^E in h_N becomes independent of N when $N \geq E$. Thus, we have shown that the quantity $N - \bar{N}_0$, that is to say, the total number of systems in states other than the lowest one is independent of N for N greater than E , but begins to depend on N as soon as N passes through this value. Thus, we have succeeded in confirming, by a purely algebraic method, the result proved above by contour integration, namely, that the anomaly sets in at N equals E .

It thus seems to be definitely established that orthodox theory cannot be relied on for any correct results in the transition region, the occupation numbers, the nature of the specific heat anomaly, and the energy content at which the transition sets in all being incorrect. One further remark about this model may be of interest. Let us make the usually correct, but not self-evident, assumption that the statistical mean state (obtained by averaging over all accessible states) is identical with the most probable state. Let us fix N_0 the occupation number of the lowest state. Then the remaining states can be occupied in $P_{N-N_0}(E)$ ways, where $P_k(E)$ is the number of ways of partitioning E into exactly k non-zero parts. The behaviour of this function

has been studied by Auluck *et al.* (1942). It is clearly an increasing function of k when k is small, and a decreasing function of k when k is becoming comparable with E . Thus, if N is below the value that makes $P_N(E)$ a maximum, we shall get the greatest number of arrangements if we make $\bar{N}_0$ zero, whereas if N is greater than this value, we shall get the largest number of arrangements if we choose $\bar{N}_0$ so that $P_{N-\bar{N}_0}(E)$ takes the maximum value. The precise behaviour of the function $P_k(E)$ is not known, but Auluck *et al.* present theoretical and numerical evidence in support of the view that there is a single maximum in the region $k \sim \frac{1}{\pi} \sqrt{\left(\frac{3E}{2}\right) \log E}$. If their conjecture proves correct, it will mean that the condensation sets in (according to the assumption that the mean state is identical with the most probable state) for a value of N differing by a factor of 2 from that predicted by the saddle-point method, and widely different from that predicted by the exact theory. It therefore seems that it is possible to construct models for which the most probable state is not identical with the mean state.

5. SOME PRELIMINARY RESULTS FOR THE 'PERFECT BOSE-EINSTEIN GAS' MODEL

There seems to be a very general relation between N the number of particles and the total energy for which a condensation sets in. In any model, let ϵ_0 be the energy-separation between the state of lowest energy and the first excited state, either or both of which may be multiple. Let $E\epsilon_0$ be the total energy. If E is less than N , there is clearly no possible arrangement of the particles for which the ground state does not contain at least one particle. If we add more particles, keeping the energy constant, we do not introduce any new possible ways in which the excited states can be occupied, the new arrangements are simply reproductions of the old ones, differing only from them by having more particles in the ground state. If the ground state is single, an exact one-to-one correspondence can be set up between the new and the old arrangements, but the correspondence is no longer exact if the ground state is multiple, as the occupation of different sub-states of the ground state does produce a few more arrangements. If the ground state is single, then the average number of particles in the excited states is quite unaffected by adding more particles, provided $N > E$, while the average increase in the number in the ground state is just the number of particles added. It is interesting to work out the energy at which degeneracy sets in in the perfect Bose-Einstein gas according to this argument. (In this model, the ground state is single.) The separation between the ground state and the first excited state, if the gas is in a cubical box of side d , is given by the expression $3\hbar^2/8md^2$, so that we have the following relation between N and E for degeneracy to set in. It is not possible to calculate the transition temperature, because we have no knowledge of the behaviour in this region of the function corresponding to $p_N(E)$ in the simple model.

$$N = E, \quad \text{with} \quad \epsilon_0 = \frac{3\hbar^2}{8md^2}. \quad (24)$$

We can, however, make a comparison with the position of the transition region as given by the orthodox theory by writing down the expressions for critical number of

atoms per unit volume as a function of temperature, and for energy content as a function of temperature and eliminating T between these two equations. These two equations are well known, see, for example Mayer & Mayer (1940, chap. 16):

$$N = \frac{(2\pi m)^{\frac{3}{2}}}{h^3} (kT)^{\frac{3}{2}} d^3 \sum_j \frac{1}{j^{\frac{3}{2}}}, \quad (25)$$

$$E\epsilon_0 = \frac{(2\pi m)^{\frac{3}{2}}}{h^3} (kT)^{\frac{3}{2}} d^3 \sum_j \frac{1}{j^{\frac{5}{2}}}. \quad (26)$$

Eliminating T between them, we find the following relation between N and E :

$$N^{\frac{2}{3}} = 0.17 E. \quad (27)$$

This differs from relation (24) by a factor of the order of $N^{\frac{2}{3}}$, so that an exact theory would give a transition beginning at a temperature far below 1°K . This investigation suggests, however, that the 'cell' model of a liquid, or some similar conception, rather than the perfect gas assumption, would, if investigated properly, be found to lead to a relationship between N and E very similar to equation (27), and may therefore be expected to lead also to a transition temperature in the right region. For suppose that each atom is supposed to be confined by the interactions with its neighbours in a volume of order of magnitude l^3 or d^3/N which is one atom's share of the total volume, then the spacing of the first excited state above the ground state is now of the order of magnitude $3h^2/8ml^2$ which differs from the perfect gas model by a factor d^2/l^2 or $N^{\frac{2}{3}}$, which is just the order of magnitude of the discrepancy between equations (24) and (27), so that the cell model, worked out properly, should give a transition temperature in the proper region.

The perfect gas model is mathematically equivalent to a certain problem in the partition of numbers into sums of squares, a subject that has been much studied by mathematicians. Unfortunately, the results hitherto available only seem to be applicable to the high temperature region, and we shall state the mathematical problem involved, in the hope that it may stimulate further work by mathematicians. Finally, we shall show the equivalence of existing theory of a perfect gas at high temperatures with certain well-known results on the representation of numbers as sums of squares. A 'representation' differs from a partition in two important respects, namely that the order of arrangement of the squares is considered, each possible order being regarded as a different representation, also the squares of negative numbers are reckoned to be different from those of positive numbers.

The mathematical problem is as follows: The energy-levels available to particles inside a cubical box of side d are given by the expression $(r^2 + s^2 + t^2) \frac{h^2}{8md^2}$, which we write $\frac{h^2}{8md^2} K(r, s, t)$ for brevity. Each of these levels may be occupied an integral number of times, or not at all. We require to know the asymptotic form for the number of distinct partitions of an integer E , representing the ratio of the total energy to the lowest possible energy-separation $h^2/8md^2$, into a sum of the numbers $K(r, s, t)$, where the order in which the K 's are arranged is ignored, but, on the other hand, K 's like $K(1, 2, 3)$, $K(2, 1, 3)$ are distinct from one another, and such K 's have

to be reckoned as 'different', in spite of the fact that they are numerically equal. In all cases, the r, s, t are positive integers (not zero).

The following result is well known, being taken from Hardy (1940),

$$S_{3N}(E) \sim \frac{\pi^{\frac{1}{2}N}}{\Gamma(\frac{3}{2}N)} E^{\frac{1}{2}N-1} \quad (28)$$

giving the number of representations of the number E as the sum of $3N$ squares. Since some of these squares may be zero, we must replace this expression, which is really the sum of the number of representations of E as $3N, 3N-1, 3N-2, \dots$, etc., non-zero squares by the expression

$$T_{3N}(E) \sim \sum_{r=0}^{\infty} \frac{\pi^{\frac{3N-r}{2}}}{\Gamma(\frac{3N-r}{2})} (-1)^r E^{\frac{3N-r}{2}-1} \quad (29)$$

which may also be written

$$T_{3N}(E) \sim S_{3N}(E) \left[1 - O\left(\frac{N^{\frac{1}{2}}}{E^{\frac{1}{2}}}\right) \right] \quad (30)$$

and the correction is thus negligible as long as $E \gg N$.

Having removed the representations involving zero squares, the next task is to remove those involving the squares of negative integers, which since the zero squares have been allowed for, simply involves dividing expression (28) by 2^{3N} . Finally, we have to allow for the fact, that, in expression (28), the same partition into squares written in a different order is counted as different. In our problem, the $3N$ squares have to be divided into N groups of three each. A permutation *within* a group of three, or an exchange of one or two squares *between* two groups of three leads, in general, to a new physical state, but a reshuffling of the *order* in which the groups of three are taken does not lead to a new physical state. Thus, the number of physically different partitions is obtained by dividing expression (28) by a further factor $N!$ giving us finally

$$Q_N(E) \sim \frac{\pi^{\frac{1}{2}N}}{\Gamma(\frac{3}{2}N)} \frac{E^{\frac{1}{2}N-1}}{N! 2^{3N}}. \quad (31)$$

The last step would be invalidated if too many of the groups of three in a typical partition were identical. I have not succeeded in making even an 'order of magnitude' estimate of the error due to this, but it seems certain on physical grounds that it becomes rapidly less important as E increases.

From expression (31), we can calculate the number of partitions which contain a state of energy $e_1 K(r, s, t)$, where e_1 is the 'energy step' $\hbar^2/8md^2$. Clearly it is the number of partitions of $E - K(r, s, t)$ into $N-1$ groups of three squares, so that the probability of finding the state (r, s, t) in a typical partition is obtained by replacing N by $N-1$ and E by $E - K(r, s, t)$ in expression (31), and then dividing this by the expression (31) as it stands, which is the total number of partitions without restriction, we obtain

$$\bar{N}_{r,s,t} \sim \frac{N(\frac{3}{2}N)^{\frac{1}{2}}}{E^{\frac{1}{2}}} \frac{8}{\pi^{\frac{1}{2}}} \left(1 - \frac{K(r, s, t)}{E} \right)^{\frac{1}{2}N-\frac{1}{2}}. \quad (32)$$

This is to be compared with the 'classical' expression

$$\bar{N}_{r,s,t} \sim N \exp \left[-\frac{\epsilon_1 K(r,s,t)}{kT} \right] / \int \int \int_0^\infty \exp \left[-\frac{(r^2+s^2+t^2)\epsilon_1}{kT} \right] dr ds dt, \quad (33)$$

and it can be seen that these two expressions agree in the limiting case of large T , provided that $E\epsilon_1 \sim \frac{3}{2}NkT$, which is the classical relation between temperature and energy content in a perfect gas.

We may thus conclude that the behaviour of the perfect gas model is qualitatively similar to that of the model with equally spaced levels, in that orthodox theory describes the situation correctly at very high temperatures but breaks down in the supposed transition region. A. R. Fraser (unpublished) has proved a rather stronger result than the one we have just obtained, namely that orthodox theory gives correct results at least down to a temperature not greater than 1.5 times the supposed transition temperature. R. B. Dingle (unpublished) has further shown that, for any model, there is a finite region of temperature starting from absolute zero, within which orthodox theory is correct.

6. CONCLUSION

The two models studied show the same general type of behaviour, namely the existence of an intermediate temperature region within which the results of orthodox theory are unreliable, the orthodox expression for the occupation numbers of the energy levels being only an approximation in this range. The investigation has confirmed the existence of the phenomenon of 'condensation into the lowest energy-level', though the perfect gas model worked out properly, would give a transition temperature far below 1° K for a perfect gas of helium atoms. On the other hand, we have seen that the 'cell' model of a liquid may well have such a transition at a temperature agreeing with the transition temperature for liquid helium, so the investigation gives us no reason to doubt either the reality of the condensation phenomenon or its relevance to the liquid helium problem.

I wish to thank Mr Dingle and Mr Fraser for helpful discussions, and for showing me their results before publication. I also wish to thank the Royal Society for awarding me the Smithson Research Fellowship during the tenure of which this work was carried out.

REFERENCES

- Auluck, F. C., Chowla, S. & Gupta, H. 1942 *J. Indian Math. Soc.* 6, 105.
 Auluck, F. C. & Kothari, D. S. 1946 *Proc. Camb. Phil. Soc.* 42, 272.
 Fowler, R. H. 1936 *Statistical mechanics*, 2nd ed. chapter 2. Cambridge Univ. Press.
 Fowler, R. H. & Jones, H. 1938 *Proc. Camb. Phil. Soc.* 34, 573.
 Hardy, G. H. 1940 *Ramanujan*, chapter 8. Cambridge Univ. Press.
 Hardy, G. H. & Ramanujan, S. 1918 *Proc. Lond. Math. Soc.* (2), 17, 75.
 Leibfried, G. 1947 *Z. Naturforsch.* 2a, 305.
 London, F. 1938 *Phys. Rev.* 54, 947.
 Mayer, J. E. & Mayer, M. G. 1940 *Statistical mechanics*, chapter 16. New York: John Wiley & Sons.
 Schubert, G. 1946 *Z. Naturforsch.* 1, 113.

An elastic theory of dislocations

By ELIZABETH H. MANN, *Cavendish Laboratory, University of Cambridge*

(Communicated by Sir Lawrence Bragg, F.R.S.—Received 22 March 1949)

A theory is presented with the intention of describing the internal stresses set up by slip processes in crystalline materials. It is based on the concept of dislocations as they occur in the mathematical theory of elasticity, and follows on the work of Volterra. The present theory starts from the assumption that only the three stress components which act across a surface in an elastic body need be continuous at that surface. This less restrictive condition on the stresses makes possible a more general type of dislocation, one with arbitrary elastic discontinuities of displacement.

The theory is stated in cylindrical polar co-ordinates, and general formulae are derived for calculating the stresses corresponding to given dislocations. Four examples are solved completely involving plane and screw dislocations in circular cylinders. The relations between the mathematical dislocations and some simple types of slip in crystals are described, and certain equivalent systems of dislocations explained.

It is shown that if the slip bands in a deformed crystal are replaced by the appropriate dislocations, the stresses in the remaining elastic region may be calculated by means of this theory.

1. INTRODUCTION

In recent years the theory of dislocations in an elastic solid has been used to explain many phenomena in crystalline materials. The fundamental notion is that slip or glide in a crystal takes place by the movement of dislocations across glide planes, the dislocation causing elastic stresses in the surrounding material. Taylor (1934) was the first to estimate these elastic stresses from Volterra's theory of dislocations, and apply the results to the work-hardening of metals. Since then dislocations have also been used in theories of the yield stress, recrystallization and thermal hardening.

It is believed that a great many dislocations may exist in a crystal after plastic deformation, and that the stresses due to them may cause considerable stored energy. If the distances between dislocations are large compared with their length, it is possible to calculate the effect of a simple two-dimensional array of dislocations by applying Volterra's theory to each and summing the results. But in many cases the dislocations appear to be crowded in narrow slip bands of high plastic deformation, separated by regions subject only to elastic stresses. Then the mathematics becomes very difficult and a new theory is required to calculate the effect of closely adjacent dislocations.

The theory to be presented in this paper may be used to calculate the stresses set up in the surrounding crystal by any system of dislocations in a slip band. These stresses are the same as if a cut were made along a glide plane, the two sides of the cut distorted elastically by different amounts and then rejoined. There is, in fact, a discontinuity of displacement across a slip band such as could be set up in this manner.

These discontinuities are more general than those described by Volterra's theory, since some of the stress components are discontinuous, as well as the displacements. This theory includes Volterra's results as particular cases.

Elastic theory cannot be applied at the centre of a Volterra dislocation, but gives the correct stresses at a short distance if the centre is excluded by an imaginary

stress-free boundary. In a similar manner the more general dislocations cause stresses in the surrounding material which can be calculated by the present theory if the actual points of discontinuity are excluded.

An outline of Volterra's theory is given in the next section, and then the new theory is described in §§ 3 to 6. For simplicity the theory is expressed in cylindrical polar co-ordinates, and examples of dislocations in cylinders are solved. The solutions could be adjusted to apply to other forms of boundary. The general screw dislocation is considered in § 7, and two examples in § 8. Finally, it is shown that one of these dislocations is equivalent to some distribution of Volterra dislocations.

2. VOLTERRA DISLOCATIONS

A Volterra (1907) dislocation is a discontinuity of displacement in an elastic body. It may be formed in any doubly-connected body, such as a closed ring, by making a cut through the ring and then displacing the two surfaces of the cut with respect to one another before rejoining them. If the surfaces of a straight cut are displaced in their own plane they can be rejoined with perfect fit, but in general a small quantity of the material must be added or removed at this stage. The relative displacement is of the 'rigid-body' type, so each side of the cut is displaced or rotated as a rigid whole.

After rejoining the sides of the cut the body is in a self-straining state, stresses existing in it even when no load is applied. To calculate such stresses for known dislocations the usual stress-strain relations and equations of equilibrium are used, but with many-valued functions for the components of displacement. By taking only one sheet of the Riemann surfaces of these functions the displacements are made discontinuous at the cut but the stresses and strains continuous.

The mathematical theory of Volterra is based on the assumption that all components of stress are continuous at the cut, and can therefore be represented by single-valued functions in the closed ring of elastic material. This assumption imposes a restriction on the displacements, since these are related to the stresses by the equations of elasticity. The components of displacement may be represented by many-valued functions, but the cyclic constants of these functions must be of a certain form. It is proved that for single-valued stresses the most general possible increment in the displacements on crossing the cut is the 'rigid body' displacement,

$$u' = a + qz - ry, \quad v' = b + rx - pz, \quad w' = c + py - qx,$$

where u', v', w' are the increments or discontinuities in the Cartesian displacements u, v, w , and a, b, c, p, q, r are constants.

3. GENERAL DISLOCATIONS

A less restrictive assumption concerning the stresses is made here. It is assumed that only the stress components which act across the surface of the cut need be continuous there; the others can be discontinuous, and all may be represented by many-valued functions.

The general dislocation is formed by making a cut and distorting the two sides of the cut elastically by applying forces to them. The sides of the cut are then rejoined

and the external forces removed, leaving an arbitrary elastic discontinuity of displacement.

A simple example is the system obtained by a uniform compression of one side of the cut, which may be rejoined to the unstrained side to give discontinuities in both displacement and direct stress components parallel to the cut. The material adjusts slightly when released, so the final state is one of uniform extension on one side and uniform compression on the other, the discontinuity remaining the same.

The displacements may be made discontinuous in any desired manner, only restricted by the equations of equilibrium and the boundary conditions.

There are new boundary conditions to be satisfied along the edges of the cut. While the cut is open the body is simply-connected, and in accordance with established theory the tractions or displacements may be prescribed at all points of the boundary surface, including the faces of the cut. If $p_{\nu i}$ ($i = 1, 2, 3$) is the component in the x_i direction ($x_i = x, y, z$) of the force acting across unit area with normal ν , these boundary conditions are either

$$p_{\nu i} = f_i \quad \text{or} \quad u_i = g_i, \quad (1)$$

where f_i and g_i are given functions of position on the surface. But when the cut is closed and set free of external forces, the following conditions must be satisfied;

$$p'_{\nu i} = 0, \quad u'_i = h_i, \quad (2)$$

where ν is the normal at any point to the surface of the cut, $p'_{\nu i}$, u'_i the increments in $p_{\nu i}$, u_i respectively on crossing the cut, and the h_i are given functions.

There is no increase in the number of boundary conditions to be satisfied, since these six equations replace a set of three equations on each side of the cut. It may readily be proved that the boundary conditions (2) with prescribed tractions or displacements over the remainder of the boundary surface are sufficient to determine uniquely the stress and strain in a doubly-connected region. The proof is similar to that of an analogous result in hydrodynamics given by Lamb (1924) under the title 'Kelvin's extension of Green's theorem'.

4. DISLOCATIONS IN CYLINDERS

The dislocations in this paper are described in cylindrical polar co-ordinates (r, θ, z) , related to Cartesian co-ordinates (x, y, z) by the equations

$$x = r \cos \theta, \quad y = r \sin \theta, \quad z = z.$$

The components of displacement are u_r, u_θ, w , the stresses $\widehat{rr}, \widehat{r\theta}$, etc., and the strains $e_{rr}, e_{r\theta}$, etc.

The equations of equilibrium are

$$\left. \begin{aligned} \frac{\partial \widehat{rr}}{\partial r} + \frac{1}{r} \frac{\partial \widehat{r\theta}}{\partial \theta} + \frac{\partial \widehat{rz}}{\partial z} + \frac{\widehat{rr} - \widehat{\theta\theta}}{r} &= 0, \\ \frac{\partial \widehat{r\theta}}{\partial r} + \frac{1}{r} \frac{\partial \widehat{\theta\theta}}{\partial \theta} + \frac{\partial \widehat{\theta z}}{\partial z} + 2 \frac{\widehat{r\theta}}{r} &= 0, \\ \frac{\partial \widehat{rz}}{\partial r} + \frac{1}{r} \frac{\partial \widehat{\theta z}}{\partial \theta} + \frac{\partial \widehat{zz}}{\partial z} + \frac{\widehat{rz}}{r} &= 0, \end{aligned} \right\} \quad (3)$$

and the stress-strain relations of the form

$$\left. \begin{aligned} \widehat{rr} &= (\lambda + 2\mu) e_{rr} + \lambda(e_{\theta\theta} + e_{zz}), \\ \widehat{\theta z} &= 2\mu e_{\theta z}, \end{aligned} \right\} \quad (4)$$

with two more pairs of similar equations, where

$$\left. \begin{aligned} e_{rr} &= \frac{\partial u_r}{\partial r}, \\ e_{\theta\theta} &= \frac{1}{r} \frac{\partial u_\theta}{\partial \theta} + \frac{u_r}{r}, \\ e_{zz} &= \frac{\partial w}{\partial z}, \\ 2e_{\theta z} &= \frac{1}{r} \frac{\partial w}{\partial \theta} + \frac{\partial u_\theta}{\partial z}, \\ 2e_{rz} &= \frac{\partial u_r}{\partial z} + \frac{\partial w}{\partial r}, \\ 2e_{r\theta} &= \frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r} + \frac{1}{r} \frac{\partial u_r}{\partial \theta}. \end{aligned} \right\} \quad (5)$$

If u'_r , u'_θ , w' represent the discontinuities in the displacements at the cut, the values of these are sufficient to specify the dislocations. If the cut is made in the half-plane $\theta = \pm \pi$, the quantities u'_r , u'_θ , w' are given functions of r and z . The simplest cases are those involving functions of r only, giving plane or 'screw' dislocations.

(i) u'_r a given function of r . The particular case in which the function is a constant is that of the dislocation shown in figure 1. The case of a linear function of r is solved as an example in § 6 and shown diagrammatically in figure 7.

(ii) u'_θ a given function of r . The function is a constant for the dislocation shown in figure 2, and a linear function of r for that in figure 3.

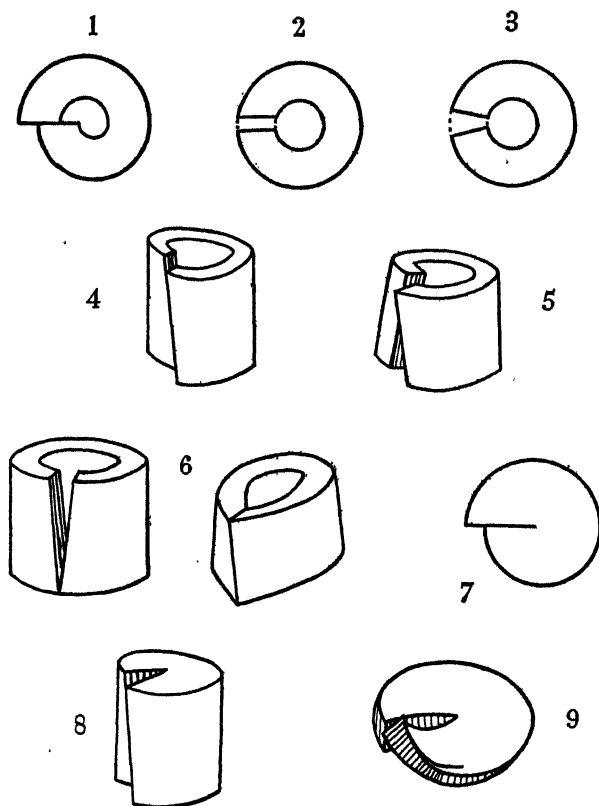
(iii) w' a given function of r . The case shown in figure 4 is the Volterra 'screw' dislocation. Figure 8 shows the dislocation corresponding to a linear function, which is considered in detail in § 8.

The general solutions of these three problems can be obtained by the use of harmonic and biharmonic functions, (i) and (ii) being cases of plane strain. Such solutions apply to cylinders of length large compared with the diameter, and in general under longitudinal forces.

From the well-known relations between plane strain and generalized plane stress these solutions can also be applied to problems of thin plates. If a radial cut is made in an annular plate and one side of the cut bent to follow a given curve before re-joining, a dislocation, such as shown in figure 9, may be set up, with discontinuities

$$w' = f(r), \quad u'_r = z \frac{df}{dr}.$$

The case of a linear function $f(r)$ is shown in figure 5 for a short cylinder. A discontinuity u'_θ linear in z gives the dislocation of figure 6. The solutions of these problems are of the form given by Love (1927*a*) and are closely related to those of the plane dislocations.



FIGURES 1 TO 9

5. THE GENERAL PLANE DISLOCATION

For plane problems, with $w = 0$ at all points, the equations of equilibrium (3) may be solved by the use of a stress function χ , satisfying the biharmonic equation

$$\left. \begin{aligned} \nabla^4 \chi &= \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right)^2 \chi = 0, \\ \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} \right)^2 \chi &= 0, \end{aligned} \right\} \quad (6)$$

at all points in the body, and related to the stress components by

$$\left. \begin{aligned} \widehat{rr} &= \frac{1}{r} \frac{\partial \chi}{\partial r} + \frac{1}{r^2} \frac{\partial^2 \chi}{\partial \theta^2}, \\ \widehat{\theta\theta} &= \frac{\partial^2 \chi}{\partial r^2}, \\ r\widehat{\theta} &= -\frac{\partial}{\partial r} \left(\frac{1}{r} \frac{\partial \chi}{\partial \theta} \right), \\ \widehat{\theta z} &= 0, \\ \widehat{rz} &= 0, \\ \widehat{zz} &= \sigma(\widehat{rr} + \widehat{\theta\theta}). \end{aligned} \right\} \quad (7)$$

The displacements are given by

$$\left. \begin{aligned} 2\mu u_r &= -\frac{\partial \chi}{\partial r} + (1-\sigma)r \frac{\partial \psi}{\partial \theta}, \\ 2\mu u_\theta &= -\frac{1}{r} \frac{\partial \chi}{\partial \theta} + (1-\sigma)r^2 \frac{\partial \psi}{\partial r}, \end{aligned} \right\} \quad (8)$$

where ψ is a harmonic function satisfying the equation

$$\frac{\partial^2}{\partial r \partial \theta} (r\psi) = \nabla^2 \chi. \quad (9)$$

Stresses and displacements may be calculated from any biharmonic function χ by equations (7), (8) and (9), and will then automatically satisfy equations (3), (4) and (5). To solve a particular problem χ must also be such as to satisfy the boundary conditions.

A set of solutions of equation (6) may be found by separating the variables, that is, assuming a solution of the form

$$\chi = R(r) \Theta(\theta) \quad (10)$$

and trying various simple forms of Θ such as $\sin n\theta$, $\cos n\theta$. The differential equations so obtained for $R(r)$ are readily solved to give a solution of the form

$$\begin{aligned} \chi &= A_0 r^2 + B_0 \log r + C_0 r^2 \log r + D_0 \\ &\quad + (Ar^2 + B \log r + Cr^2 \log r + D) \theta \\ &\quad + \left(A_1 r + \frac{B_1}{r} + C_1 r^3 + D_1 r \log r \right) \sin \theta + (Er + Fr \log r) \theta \sin \theta \\ &\quad + \left(A'_1 r + \frac{B'_1}{r} + C'_1 r^3 + D'_1 r \log r \right) \cos \theta + (E'r + F'r \log r) \theta \cos \theta \\ &\quad + \sum_{n=2}^{\infty} \left(A_n r^n + \frac{B_n}{r^n} + C_n r^{n+2} + \frac{D_n}{r^{n-2}} \right) \sin n\theta \\ &\quad + \sum_{n=2}^{\infty} \left(A'_n r^n + \frac{B'_n}{r^n} + C'_n r^{n+2} + \frac{D'_n}{r^{n-2}} \right) \cos n\theta. \end{aligned} \quad (11)$$

Part of this solution was given by Michell (1899). The terms with coefficients C_0 , D_1 , D'_1 , E and E' correspond to Volterra dislocations, those with coefficients A , B , C , F and F' to more general dislocations with discontinuous stresses.

Many of the solutions used in this paper do not satisfy the condition (10) so cannot be found by this method. It is more convenient to use the complex variable

$$z = x + iy = r e^{i\theta},$$

and the conjugate complex $\bar{z} = x - iy = r e^{-i\theta}$,

and write (6) as
$$\frac{\partial^4 \chi}{\partial z^2 \partial \bar{z}^2} = 0.$$

The general real solution of this equation is

$$\chi = \Re \{ \bar{z} \phi_1(z) + \phi_2(z) \},$$

where $\Re$ denotes 'the real part of'. By choosing ϕ_1 and ϕ_2 to contain $\log z$ multiplied by some power of z the following solutions are obtained:

$$\begin{aligned} \chi_m &= r^m \{ P_m (\log r \cos m\theta - \theta \sin m\theta) + Q_m (\log r \sin m\theta + \theta \cos m\theta) \\ &\quad + R_m (\log r \cos \overline{m-2}\theta - \theta \sin \overline{m-2}\theta) + S_m (\log r \sin \overline{m-2}\theta + \theta \cos \overline{m-2}\theta) \}, \end{aligned} \quad (12)$$

for integral values of m , and arbitrary constants P_m , Q_m , R_m and S_m . It is seen that these do not in general satisfy (10) except certain terms in the cases $m = 0, 1, 2$ which appeared in (11).

The terms with coefficients P_m , Q_m are harmonic, and those in R_m , S_m are of the form

$$r^2\phi,$$

where ϕ is harmonic. Only the latter terms make any contribution to ψ in (9), which gives

$$\begin{aligned}\psi &= 4 \int \phi d\theta \\ &= \frac{4r^{m-2}}{m-2} \left\{ R_m \left[\log r \sin \overline{m-2}\theta + \theta \cos \overline{m-2}\theta - \frac{\sin \overline{m-2}\theta}{(m-2)} \right] \right. \\ &\quad \left. + S_m \left[-\log r \cos \overline{m-2}\theta + \theta \sin \overline{m-2}\theta + \frac{\cos \overline{m-2}\theta}{(m-2)} \right] \right\},\end{aligned}$$

$$\text{when } m \neq 2, \text{ and } \psi = 4R_2\theta \log r + 2S_2(\theta^2 - (\log r)^2), \quad (13)$$

when $m = 2$. The displacements corresponding to (12) are given by

$$\left. \begin{aligned}2\mu u_r &= r^{m-1} \{ P_m [-(m \log r + 1) \cos m\theta + m\theta \sin m\theta] \\ &\quad Q_m [-(m \log r + 1) \sin m\theta - m\theta \cos m\theta] \\ &\quad R_m [(-(m-4+4\sigma) \log r - 1) \cos \overline{m-2}\theta + (m-4+4\sigma) \theta \sin \overline{m-2}\theta] \\ &\quad S_m [(-(m-4+4\sigma) \log r - 1) \sin \overline{m-2}\theta - (m-4+4\sigma) \theta \cos \overline{m-2}\theta] \}, \\ 2\mu u_\theta &= r^{m-1} \{ P_m [(m \log r + 1) \sin m\theta + m\theta \cos m\theta] \\ &\quad Q_m [-(m \log r + 1) \cos m\theta + m\theta \sin m\theta] \\ &\quad R_m [(m+2-4\sigma) \log r + 1) \sin \overline{m-2}\theta + (m+2-4\sigma) \theta \cos \overline{m-2}\theta] \\ &\quad S_m [-(m+2-4\sigma) \log r - 1) \cos \overline{m-2}\theta + (m+2-4\sigma) \theta \sin \overline{m-2}\theta] \}. \end{aligned} \right\} \quad (14)$$

The stress components corresponding to χ_m in (12) are

$$\left. \begin{aligned}\widehat{rr} &= r^{m-2} \{ P_m [(-m(m-1) \log r - 2m + 1) \cos m\theta + m(m-1) \theta \sin m\theta] \\ &\quad Q_m [(-m(m-1) \log r - 2m + 1) \sin m\theta - m(m-1) \theta \cos m\theta] \\ &\quad R_m [(-(m-1)(m-4) \log r - 2m + 5) \cos \overline{m-2}\theta + (m-1)(m-4) \theta \sin \overline{m-2}\theta] \\ &\quad S_m [(-(m-1)(m-4) \log r - 2m + 5) \sin \overline{m-2}\theta - (m-1)(m-4) \theta \cos \overline{m-2}\theta] \}, \\ \widehat{\theta\theta} &= r^{m-2} \{ P_m [(m(m-1) \log r + 2m - 1) \cos m\theta - m(m-1) \theta \sin m\theta] \\ &\quad Q_m [(m(m-1) \log r + 2m - 1) \sin m\theta + m(m-1) \theta \cos m\theta] \\ &\quad R_m [(m(m-1) \log r + 2m - 1) \cos \overline{m-2}\theta - m(m-1) \theta \sin \overline{m-2}\theta] \\ &\quad S_m [(m(m-1) \log r + 2m - 1) \sin \overline{m-2}\theta + m(m-1) \theta \cos \overline{m-2}\theta] \}, \\ \widehat{r\theta} &= r^{m-2} \{ P_m [(m(m-1) \log r + 2m - 1) \sin m\theta + m(m-1) \theta \cos m\theta] \\ &\quad Q_m [(-m(m-1) \log r - 2m + 1) \cos m\theta + m(m-1) \theta \sin m\theta] \\ &\quad R_m [(m-2)(m-1) \log r + 2m - 3) \sin \overline{m-2}\theta + (m-2)(m-1) \theta \cos \overline{m-2}\theta] \\ &\quad S_m [-(m-2)(m-1) \log r - 2m + 3) \cos \overline{m-2}\theta + (m-2)(m-1) \theta \sin \overline{m-2}\theta] \}. \end{aligned} \right\} \quad (15)$$

The problem is to find the stresses in a cylinder due to the general dislocation of type (i)

$$u'_r = f(r),$$

and of type (ii)

$$u'_\theta = g(r).$$

Stress functions such as (11) and (12) may be used to satisfy the equations of equilibrium in $b \leq r \leq a$, $-\pi \leq \theta \leq \pi$ and the arbitrary constants are to be chosen so that

$$\bar{r}\bar{r} = 0 = \bar{r}\bar{\theta}$$

on $r = a$, $r = b$, besides satisfying conditions (2) at the cut.

From (14) the discontinuities in u_r and u_θ are given by

$$2\mu u'_r = 2\pi(-1)^{m-1} r^{m-1} (mQ_m + (m-4+4\sigma)S_m),$$

$$2\mu u'_\theta = 2\pi(-1)^m r^{m-1} (mP_m + (m+2-4\sigma)R_m).$$

If $f(r)$ and $g(r)$ are represented by series of powers of r

$$f(r) = \Sigma a_m r^m, \quad g(r) = \Sigma b_m r^m,$$

with satisfactory convergence in $b \leq r \leq a$, the required discontinuities can be set up by choosing

$$\left. \begin{aligned} 2\pi(-1)^{m-1} (mQ_m + (m-4+4\sigma)S_m) &= 2\mu a_{m-1}, \\ 2\pi(-1)^m (mP_m + (m+2-4\sigma)R_m) &= 2\mu b_{m-1}. \end{aligned} \right\} \quad (16)$$

Two more equations are required to determine completely the constants P_m , Q_m , R_m and S_m , and these are provided by the other condition (2),

$$\bar{\theta}\bar{\theta}' = 0 = \bar{r}\bar{\theta}'$$

at $\theta = \pm\pi$. From (15)

$$\left. \begin{aligned} \bar{\theta}\bar{\theta}' &= 2\pi(-1)^m r^{m-2} m(m-1) (Q_m + S_m) = 0, \\ \bar{r}\bar{\theta}' &= 2\pi(-1)^m r^{m-2} (m(m-1)P_m + (m-1)(m-2)R_m) = 0. \end{aligned} \right\} \quad (17)$$

The constants Q_m and S_m appear in u'_r and $\bar{\theta}\bar{\theta}'$ only, and P_m and R_m in u'_θ and $\bar{r}\bar{\theta}'$. The equations are sufficient to determine the coefficients in the part of the stress function of the form (12). Further terms must be added to satisfy the other boundary conditions

$$\bar{r}\bar{r} = 0 = \bar{r}\bar{\theta}$$

on $r = a$ and $r = b$, and this can be done with single-valued terms from (11)

$$\begin{aligned} \chi &= A_0 r^2 + \left(A_1 r + \frac{B_1}{r} + C_1 r^3 \right) \sin \theta + \left(A'_1 r + \frac{B'_1}{r} + C'_1 r^3 \right) \cos \theta \\ &+ \sum_{n=2}^{\infty} \left(A_n r^n + \frac{B_n}{r^n} + C_n r^{n+2} + \frac{D_n}{r^{n-2}} \right) \sin n\theta \\ &+ \sum_{n=2}^{\infty} \left(A'_n r^n + \frac{B'_n}{r^n} + C'_n r^{n+2} + \frac{D'_n}{r^{n-2}} \right) \cos n\theta. \end{aligned} \quad (18)$$

These terms give the stresses

$$\left. \begin{aligned} \widehat{r\tau} &= 2A_0 + \left(-\frac{2B_1}{r^3} + 2C_1 r \right) \sin \theta + \left(-\frac{2B'_1}{r^3} + 2C'_1 r \right) \cos \theta \\ &\quad - \sum_{n=2}^{\infty} \left(n(n-1) A_n r^{n-2} + \frac{n(n+1) B_n}{r^{n+2}} + (n-2)(n+1) C_n r^n + (n+2)(n-1) \frac{D_n}{r^n} \right) \sin n\theta \\ &\quad - \sum_{n=2}^{\infty} \left(n(n-1) A'_n r^{n-2} + \frac{n(n+1) B'_n}{r^{n+2}} + (n-2)(n+1) C'_n r^n + (n+2)(n-1) \frac{D'_n}{r^n} \right) \cos n\theta, \\ \widehat{r\theta} &= \left(\frac{2B_1}{r^3} - 2C_1 r \right) \cos \theta - \left(\frac{2B'_1}{r^3} - 2C'_1 r \right) \sin \theta \\ &\quad - \sum_{n=2}^{\infty} n \left((n-1) A_n r^{n-2} - \frac{(n+1) B_n}{r^{n+2}} + (n+1) C_n r^n - \frac{(n-1) D_n}{r^n} \right) \cos n\theta \\ &\quad + \sum_{n=2}^{\infty} n \left((n-1) A'_n r^{n-2} - \frac{(n+1) B'_n}{r^{n+2}} + (n+1) C'_n r^n - \frac{(n-1) D'_n}{r^n} \right) \sin n\theta, \end{aligned} \right\} \quad (19)$$

where $r = a$ or b . The series cancel the boundary stresses of the form (15) if A_n , B_n , C_n , D_n , A'_n , B'_n , C'_n and D'_n are suitably chosen. The right-hand sides of equations (19) are made equal and opposite to the Fourier expansions of $\widehat{r\tau}$ and $\widehat{r\theta}$ from (15) in $-\pi < \theta < \pi$. The corresponding process for one boundary only is described in detail by Bickley (1928). The more general case of two boundaries is solved in a similar manner, the equations proving sufficient to determine all the constants, as demonstrated by the following examples.

6. EXAMPLES OF PLANE DISLOCATIONS

(a) $u'_\theta = \alpha$ (figure 2)

This is a Volterra dislocation since each side of the cut is displaced as a rigid whole. The required discontinuity is provided by a stress function of the form (12), in which

$$m = 1, \quad Q_m = 0 = S_m.$$

Then by (16)

$$-2\pi(P_1 + (3-4\sigma)R_1) = 2\mu\alpha,$$

and $\widehat{r\theta}$ and $\widehat{\theta\theta}$ are continuous across the cut because (17) has $(m-1)$ as a factor.

The stress function has the form

$$\chi = (P_1 + R_1)r \log r \cos \theta - (P_1 - R_1)r\theta \sin \theta + \left(\frac{B'_1}{r} + C'_1 r^3 \right) \cos \theta,$$

with surface stresses

$$\widehat{r\tau} = \left(-\frac{P_1}{r} + \frac{3R_1}{r} - \frac{2B'_1}{r^3} + 2C'_1 r \right) \cos \theta,$$

$$\widehat{r\theta} = \left(\frac{P_1}{r} + \frac{S_1}{r} - \frac{2B'_1}{r^3} + 2C'_1 r \right) \sin \theta.$$

These are zero when r takes the values a and b if

$$P_1 = -\frac{\mu\alpha}{4\pi(1-\sigma)} = R_1,$$

$$B'_1 = P_1 a^2 b^2 / (a^2 + b^2),$$

$$C'_1 = -P_1 / (a^2 + b^2).$$

The solution of the problem is therefore given by the stress function

$$\chi = -\frac{\mu\alpha}{4\pi(1-\sigma)} \left[2r \log r + \frac{a^2 b^2}{a^2 + b^2} \frac{1}{r} - \frac{r^3}{a^2 + b^2} \right] \cos \theta.$$

(b) $u'_r = \alpha r$ (figure 7)

Let

$$P_m = 0 = R_m$$

in (12) for all m , and

$$Q_m = 0 = S_m \quad (m \neq 2).$$

Equations (16) and (17) give

$$Q_2 = -\frac{\mu\alpha}{4\pi(1-\sigma)} = -T_2.$$

The stress function is of the form

$$\begin{aligned} \chi = & -\frac{\mu\alpha}{4\pi(1-\sigma)} (r^2 \log r \sin 2\theta + r^2 \theta \cos 2\theta - r^2 \theta) + \left(\frac{B_1}{r} + C_1 r^3 \right) \sin \theta \\ & + \sum_{n=2}^{\infty} \left(A_n r^n + \frac{B_n}{r^n} + C_n r^{n+2} + \frac{D_n}{r^{n-2}} \right) \sin n\theta, \end{aligned}$$

and the surface stresses are

$$\begin{aligned} \widehat{rr} = & \frac{\mu\alpha}{4\pi(1-\sigma)} [(2 \log r + 3) \sin 2\theta + 2\theta(1 + \cos 2\theta)] - 2 \left(\frac{B_1}{r^3} - C_1 r \right) \sin \theta \\ & - \sum_{n=2}^{\infty} \left(n(n-1) A_n r^{n-2} + n(n+1) \frac{B_n}{r^{n+2}} + (n-2)(n+1) C_n r^n \right. \\ & \left. + (n+2)(n-1) \frac{D_n}{r^n} \right) \sin n\theta, \end{aligned}$$

$$\begin{aligned} \widehat{r\theta} = & \frac{\mu\alpha}{4\pi(1-\sigma)} [(2 \log r + 3) \cos 2\theta - 1 - 2\theta \sin 2\theta] + 2 \left(\frac{B_1}{r^3} - C_1 r \right) \cos \theta \\ & - \sum_{n=2}^{\infty} n \left((n-1) A_n r^{n-2} - (n+1) \frac{B_n}{r^{n+2}} + (n+1) C_n r^n - (n-1) \frac{D_n}{r^n} \right) \cos n\theta, \end{aligned}$$

in which r takes the values a or b . To make these equal to zero, the functions

$$\frac{\mu\alpha}{2\pi(1-\sigma)} \theta (\cos 2\theta + 1) \quad \text{and} \quad \frac{\mu\alpha}{2\pi(1-\sigma)} \theta \sin 2\theta$$

are expressed as series of sines and cosines respectively in the range $-\pi < \theta < \pi$:

$$\theta (\cos 2\theta + 1) = \sum_{n=1}^{\infty} a_n \sin n\theta,$$

where

$$a_n = 4(-1)^{n+1} (n^2 - 2) / n(n^2 - 4) \quad (n \neq 2),$$

$$a_2 = -\frac{5}{4},$$

and

$$\theta \sin 2\theta = \frac{1}{2} b_0 + \sum_{n=1}^{\infty} b_n \cos n\theta,$$

where

$$b_n = 4(-1)^n / (n^2 - 4) \quad (n \neq 2),$$

$$b_2 = -\frac{1}{4}.$$

In particular $b_0 = -1$, which cancels the constant term of $\bar{r}\bar{\theta}$. The surface stresses vanish if, on substituting these series in $\bar{r}\bar{r}$ and $\bar{r}\bar{\theta}$, the coefficient of each $\cos n\theta$ or $\sin n\theta$ vanishes at $r = a$ and $r = b$. This gives the following equations:

$$\frac{2B_1}{a^3} - 2C_1 a = \frac{2\mu\alpha}{3\pi(1-\sigma)},$$

$$2A_2 + \frac{6B_2}{a^4} + \frac{4D_2}{a^2} = \frac{\mu\alpha}{4\pi(1-\sigma)} (2\log a + \frac{1}{2}),$$

$$2A_2 - \frac{6B_2}{a^4} + 6C_2 a^2 - \frac{2D_2}{a^2} = \frac{\mu\alpha}{4\pi(1-\sigma)} (2\log a + \frac{7}{2}),$$

$$n(n-1)A_n a^{n-2} + n(n+1)\frac{B_n}{a^{n+2}} + (n-2)(n+1)C_n a^n + (n+2)(n-1)\frac{D_n}{a^2} = \frac{2\mu\alpha(-1)^{n+1}(n^2-2)}{\pi(1-\sigma)n(n^2-4)},$$

$$n\left((n-1)A_n a^{n-2} - (n+1)\frac{B_n}{a^{n+2}} + (n+1)C_n a^n - (n-1)\frac{D_n}{a^n}\right) = \frac{2\mu\alpha(-1)^{n+1}}{\pi(1-\sigma)(n^2-4)},$$

and the same equations with a replaced by b . From these equations all the arbitrary constants may be determined for any values of a and b .

The particular case $b = 0$ is that shown in figure 7; the displacements are zero at the axis of the cylinder but the stress components are singular there. The coefficients B_n and D_n are all zero in this case, and the other constants have the values

$$C_1 = -\mu\alpha/3\pi(1-\sigma)a,$$

$$A_2 = \mu\alpha(\log a + \frac{1}{4})/4\pi(1-\sigma),$$

$$C_2 = \mu\alpha/8\pi(1-\sigma)a^2,$$

$$A_n = -\mu\alpha(-1)^n/\pi(1-\sigma)(n^2-4)a^{n-2},$$

$$C_n = \mu\alpha(-1)^n/\pi(1-\sigma)n(n+2)a^n.$$

The final stress components are then

$$\bar{r}\bar{r} = \frac{\mu\alpha}{4\pi(1-\sigma)} \left\{ -\frac{8r}{3a} \sin \theta + 2\theta(1 + \cos 2\theta) + \left(2\log \frac{r}{a} + \frac{5}{2}\right) \sin 2\theta + 4 \sum_{n=3}^{\infty} (-1)^n \left[\frac{n(n-1)}{n^2-4} \left(\frac{r}{a}\right)^{n-2} - \frac{(n-2)(n+1)}{n(n+2)} \left(\frac{r}{a}\right)^n \right] \sin n\theta \right\},$$

$$\bar{\theta}\bar{\theta} = \frac{\mu\alpha}{4\pi(1-\sigma)} \left\{ -\frac{8r}{a} \sin \theta + 2\theta(1 - \cos 2\theta) - \left(2\log \frac{r}{a} + \frac{5}{2} - \frac{6r^2}{a^2}\right) \sin 2\theta - 4 \sum_{n=3}^{\infty} (-1)^n \left[\frac{n(n-1)}{n^2-4} \left(\frac{r}{a}\right)^{n-2} - \frac{(n+1)}{n} \left(\frac{r}{a}\right)^n \right] \sin n\theta \right\},$$

$$\bar{r}\bar{\theta} = \frac{\mu\alpha}{4\pi(1-\sigma)} \left\{ \frac{8r}{3a} \cos \theta - 1 - 2\theta \sin 2\theta + \left(2\log \frac{r}{a} + \frac{5}{2} - \frac{3r^2}{a^2}\right) \cos 2\theta + 4 \sum_{n=3}^{\infty} (-1)^n \left[\frac{n(n-1)}{n^2-4} \left(\frac{r}{a}\right)^{n-2} - \frac{n+1}{n+2} \left(\frac{r}{a}\right)^n \right] \cos n\theta \right\},$$

giving a logarithmic singularity at the z -axis, and a discontinuity in $\bar{r}\bar{r}$ at the cut.

The displacements are given by

$$2\mu u_r = \frac{\mu \alpha r}{4\pi(1-\sigma)} [2\theta(1-2\sigma + \cos 2\theta) + (2 \log r + 1) \sin 2\theta + \Sigma \dots],$$

$$2\mu u_\theta = \frac{\mu \alpha r}{4\pi(1-\sigma)} [-4(1-\sigma) \log r - 1 - 2\theta \sin 2\theta + (2 \log r + 1) \cos 2\theta + \Sigma \dots],$$

where $\Sigma \dots$ indicates single-valued terms involving only positive powers of r . The displacements are therefore zero or constant at the z -axis, and may be made zero by a rigid displacement of the whole body.

7. THE GENERAL SCREW DISLOCATION

A screw dislocation is here defined as any discontinuity of w which is a function of r only, independent of z .

Consider equations (3), (4) and (5) for w independent of z . The only stresses then affected by w are

$$\widehat{\theta z} = \frac{\mu}{r} \frac{\partial w}{\partial \theta}, \quad \widehat{r z} = \mu \frac{\partial w}{\partial r},$$

and the equations of equilibrium are satisfied if

$$\frac{\partial \widehat{r z}}{\partial r} + \frac{1}{r} \frac{\partial \widehat{\theta z}}{\partial \theta} + \frac{\widehat{r z}}{r} = 0,$$

i.e. if

$$\nabla^2 w = 0.$$

It is therefore possible to satisfy the equations of equilibrium by a displacement w alone, provided w is a plane harmonic function. If w were not harmonic, but bi-harmonic, certain displacements u_r and u_θ would be required to maintain equilibrium. For the dislocations of this section harmonic forms of w are sufficient.

Let the required discontinuity of w have the form

$$\begin{aligned} w' &= f(r) \\ &= \Sigma C_n r^n \quad (\text{in } b \leq r \leq a). \end{aligned}$$

Consider only $w' = C_m r^m$ and later sum the solutions if desired. As shown in § 5 a discontinuity of order m may come from the real or imaginary part of

$$z^m \log z \quad (z = x + iy),$$

either of which is a plane harmonic function. To give a discontinuity at $\theta = \pm \pi$ take the imaginary part

$$w_m = K r^m (\log r \sin m\theta + \theta \cos m\theta),$$

which has the discontinuity $C_m r^m$ if

$$K = C_m (-1)^m / 2\pi.$$

Further harmonic terms may be added in the form of sine or cosine series. The sine series is sufficient here since w is odd in θ :

$$w = K r^m (\log r \sin m\theta + \theta \cos m\theta) + \sum_{n=1}^{\infty} (A_n r^n + B_n r^{-n}) \sin n\theta. \quad (20)$$

This w satisfies the equations of equilibrium and causes the following stresses

$$\widehat{rr} = 0 = \widehat{\theta\theta} = \widehat{zz} = \widehat{r\theta},$$

$$\widehat{rz} = \mu K r^{m-1} [(m \log r + 1) \sin m\theta + m\theta \cos m\theta] + \mu \sum_{n=1}^{\infty} n (A_n r^{n-1} - B_n r^{-n-1}) \sin n\theta,$$

$$\widehat{\theta z} = \mu K r^{m-1} [(m \log r + 1) \cos m\theta - m\theta \sin m\theta] + \mu \sum_{n=1}^{\infty} n (A_n r^{n-1} + B_n r^{-n-1}) \cos n\theta.$$

The surface stress $\widehat{rz}$ on $r = a$, $r = b$ may be removed by suitable choice of the coefficients A_n and B_n . Their values are determined by expressing $m\theta \cos m\theta$ as a sine series

$$\theta \cos m\theta = \sum_{n=1}^{\infty} a_n \sin n\theta,$$

where the a_n are Fourier coefficients

$$a_n = (-1)^{m+n+1} 2n / (n^2 - m^2) \quad (n \neq m),$$

$$a_m = -1/2m.$$

Then $\widehat{rz}$ is zero at $r = a$, $r = b$ if

$$\left. \begin{aligned} A_n a^n - B_n a^{-n} &= K a^m (-1)^{m+n} 2m / (n^2 - m^2) \\ A_n b^n - B_n b^{-n} &= K b^m (-1)^{m+n} 2m / (n^2 - m^2) \end{aligned} \right\} \quad (n \neq m),$$

$$A_m a^m - B_m a^{-m} = -K a^m (m \log a + \frac{1}{2}) / m,$$

$$A_m b^m - B_m b^{-m} = -K b^m (m \log b + \frac{1}{2}) / m,$$

giving the values

$$\left. \begin{aligned} A_n &= 2K (-1)^{m+n} \frac{m}{(n^2 - m^2)} \frac{(a^{m+n} - b^{m+n})}{(a^{2n} - b^{2n})} \\ B_n &= 2K (-1)^{m+n} \frac{m}{(n^2 - m^2)} \frac{(a^{m+n} b^{2n} - a^{2n} b^{m+n})}{(a^{2n} - b^{2n})} \end{aligned} \right\} \quad (n \neq m),$$

$$A_m = -K \left(\frac{(a^{2m} \log a - b^{2m} \log b)}{(a^{2m} - b^{2m})} + \frac{1}{2m} \right),$$

$$B_m = -K \frac{(\log a - \log b) a^{2m} b^{2m}}{(a^{2m} - b^{2m})}.$$

With these values of A_n , B_n it is found that the tractions acting across any plane $z = \text{constant}$ reduce to a couple. For the force resultants vanish leaving only a couple of moment M

$$\begin{aligned} M &= \int_b^a \int_{-\pi}^{\pi} \widehat{\theta z} r^2 d\theta dr \\ &= -\mu K m \int_b^a \int_{-\pi}^{\pi} \theta \sin m\theta r^{m+1} d\theta dr \\ &= (-1)^m 2\pi \mu K (a^{m+2} - b^{m+2}) / (m+2). \end{aligned}$$

If, therefore, the ends of the cylinder remain undisplaced, a torsional couple must be applied to them by the surrounding medium. If this couple is not applied the ends

rotate, so that the cut no longer lies along a generator of the cylinder, but curves slightly about the axis, as shown in figures 4 and 8.

Mathematically the couple is removed by adding a simple torsion solution of the form

$$u_\theta = -\tau rz, \quad u_r = 0 = w.$$

The stresses from this are zero except

$$\widehat{\theta z} = -\mu \tau r,$$

giving a resultant couple of moment M' :

$$\begin{aligned} M' &= \mu \tau 2\pi \int r^3 dr \\ &= \frac{1}{2} \mu \tau \pi (a^4 - b^4). \end{aligned}$$

The moments M and M' are equal and opposite if

$$\tau = 4(-1)^m K(a^{m+2} - b^{m+2}) / (m+2) (a^4 - b^4).$$

The remaining surface stresses consist of shears $\widehat{rz}$ and $\widehat{\theta z}$ on the ends of the cylinder such that the resultant forces and moments on each end are zero. Love (1927*b*) has shown that the stresses due to such shears may be expressed in terms of Bessel functions, and that they decrease exponentially with increasing distance from the ends. If the length of the cylinder is large compared with its diameter the stresses over most of the body are accurately obtained by neglecting the local perturbations at the ends. The solution is then given by

$$\begin{aligned} u_r &= 0, \\ u_\theta &= 4(-1)^{m+1} K(a^{m+2} - b^{m+2}) rz / (m+2) (a^4 - b^4), \\ w &= Kr^m (\log r \sin m\theta + \theta \cos m\theta) + \sum_{n=1}^{\infty} (A_n r^n + B_n r^{-n}) \sin n\theta, \end{aligned}$$

where A_n, B_n have the values given above.

8. EXAMPLES

(a) $w' = \text{constant} = \alpha$ (figure 4)

This is the simple or Volterra screw dislocation, which does not involve a discontinuity of stress.

Let $m = 0$ in (20) so that the required form is

$$w = \frac{\alpha \theta}{2\pi} + \sum_{n=1}^{\infty} (A_n r^n + B_n r^{-n}) \sin n\theta.$$

The stress $\widehat{rz}$ is given by

$$\widehat{rz} = \sum_{n=1}^{\infty} n(A_n r^{n-1} - B_n r^{-n-1}) \sin n\theta,$$

which vanishes on $r = a, r = b$ if all A_n and B_n are zero.

The other stress component is

$$\widehat{\theta z} = \frac{\mu \alpha}{2\pi r},$$

with resultant couple on the ends $z = \text{constant}$ of moment

$$\int_b^a 2\pi r^2 \widehat{\theta z} dr = \frac{1}{2} \mu \alpha (a^2 - b^2).$$

If the ends are free, this couple is removed by adding an equal and opposite one by means of the solution

$$u_\theta = -\tau rz, \quad \widehat{\theta z} = -\mu \tau r,$$

which causes a resultant couple with moment

$$\int 2\pi r^2 (-\mu \tau r) dr = -\frac{1}{2} \mu \pi \tau (a^4 - b^4),$$

on the ends of the cylinder. This cancels the first couple if

$$\tau = \frac{\alpha}{\pi(a^2 + b^2)}.$$

The remaining stresses at the ends are

$$\widehat{\theta z} = \frac{\mu \alpha}{2\pi r} \left(1 - \frac{2r^2}{a^2 + b^2} \right),$$

which have only a local effect on a long cylinder, and the elastic state due to the screw dislocation is given by

$$u_r = 0, \quad u_\theta = -\alpha rz / \pi(a^2 + b^2), \quad w = \alpha \theta / 2\pi.$$

(b) $w' = \alpha r$ (figure 8)

Let $m = 1$ in (20), so that w has the form

$$w = -\frac{\alpha r}{2\pi} (\log r \sin \theta + \theta \cos \theta) + \sum_{n=1}^{\infty} (A_n r^n + B_n r^{-n}) \sin n\theta.$$

The stress $\widehat{rz}$ from this w is discontinuous at the cut,

$$\widehat{rz} = -\frac{\mu \alpha}{2\pi} \left((\log r + 1) \sin \theta + \theta \cos \theta \right) + \mu \sum_{n=1}^{\infty} n (A_n r^{n-1} - B_n r^{-n-1}) \sin n\theta,$$

and is to be made zero on $r = a$, $r = b$ by suitable choice of A_n , B_n . The function $\theta \cos \theta$ may be represented by a Fourier series in the range $-\pi < \theta < \pi$

$$\theta \cos \theta = \sum_{n=1}^{\infty} a_n \sin n\theta,$$

where

$$a_n = (-1)^n 2n / (n^2 - 1) \quad (n > 1),$$

$$a_1 = -\frac{1}{2}.$$

Then $\widehat{rz}$ vanishes at $r = a$, $r = b$ if

$$\left. \begin{aligned} A_n &= \frac{\alpha (-1)^n (a^{n+1} - b^{n+1})}{\pi (n^2 - 1) (a^{2n} - b^{2n})} \\ B_n &= -\frac{\alpha (-1)^n (a^{2n} b^{n+1} - a^{n+1} b^{2n})}{\pi (n^2 - 1) (a^{2n} - b^{2n})} \end{aligned} \right\} \quad (n > 1),$$

$$A_1 = \frac{\alpha}{2\pi} \left(\frac{a^2 \log a - b^2 \log b}{(a^2 - b^2)} + \frac{1}{2} \right),$$

$$B_1 = \frac{\alpha}{2\pi} \frac{a^2 b^2 (\log a - \log b)}{(a^2 - b^2)}.$$

The other shear stress $\widehat{\theta z}$ has the value

$$\widehat{\theta z} = -\frac{\mu\alpha}{2\pi}((\log r + 1)\cos\theta - \theta\sin\theta) + \mu \sum_{n=1}^{\infty} n(A_n r^{n-1} + B_n r^{-n-1})\cos n\theta,$$

with a resultant couple on the z -plane of moment

$$\begin{aligned} \int_b^a \int_{-\pi}^{\pi} \widehat{\theta z} r^2 d\theta dr &= \frac{\mu\alpha}{6\pi}(a^3 - b^3) \int \theta \sin\theta d\theta \\ &= \frac{\mu\alpha(a^3 - b^3)}{3}. \end{aligned}$$

To free the ends of the cylinder a simple torsion solution is added as before:

$$u_{\theta} = -\tau rz,$$

with resultant couple

$$M = -\frac{1}{2}\mu\pi\tau(a^4 - b^4).$$

Therefore τ is chosen to have the value

$$\tau = \frac{2\alpha(a^3 - b^3)}{3\pi(a^4 - b^4)}.$$

The boundary surfaces are then free except for shear stresses on the ends which have zero resultant forces and moment. Their effect is negligible at a distance from the ends and the following solution is valid:

$$\begin{aligned} u_r &= 0, \\ u_{\theta} &= -\frac{2\alpha(a^3 - b^3)rz}{3\pi(a^4 - b^4)}, \\ w &= -\frac{\alpha r}{2\pi}(\log r \sin\theta + \theta \cos\theta) + \sum_{n=1}^{\infty} (A_n r^n + B_n r^{-n})\sin n\theta, \end{aligned}$$

where A_n, B_n have the values given above.

If $b \rightarrow 0$, the displacements remain finite to the z -axis but the stresses have a logarithmic singularity there. The coefficients A_n and B_n approach limiting values,

$$\left. \begin{aligned} A_1 &\rightarrow \frac{\alpha}{4\pi}(1 + 2\log a), \\ B_1 &\rightarrow 0, \\ A_n &\rightarrow \frac{\alpha(-1)^n}{\pi a^{n-1}(n^2 - 1)} \\ B_n &\rightarrow 0 \end{aligned} \right\} \quad (n > 1),$$

and the corresponding solution is

$$\begin{aligned} u_r &= 0, \\ u_{\theta} &= \frac{2\alpha rz}{3\pi a}, \\ w &= -\frac{\alpha r}{2\pi} \left[\left(\log \frac{r}{a} - \frac{1}{2} \right) \sin\theta + \theta \cos\theta \right] + \frac{\alpha a}{\pi} \sum_{n=2}^{\infty} \frac{(-1)^n}{(n^2 - 1)} \left(\frac{r}{a} \right)^n \sin n\theta. \end{aligned}$$

9. EQUIVALENT DISLOCATIONS

In crystal physics a dislocation usually means a disturbed region in the crystal lattice, forming a boundary between areas of a slip plane in which slip has and has not taken place. The disturbed state of the lattice is represented diagrammatically in figure 10*a*, the slip plane being normal to the paper. It may be supposed that glide has taken place on the right of the encircled area, and not on the left, or vice versa. The lattice on either side has rejoined perfectly, so that only by examining the boundary of the crystal can it be decided whether the deformation is as in figure 10*b* or 10*c*.

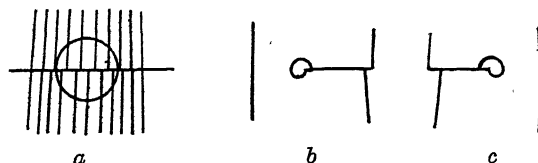


FIGURE 10

The stresses around such a region are those of the first Volterra dislocation $u_r' = \alpha$ of figure 1, which is known as a 'line' dislocation.

The same singularity in the lattice could have been set up by removing one vertical row of atoms from the centre of the upper half plane in figure 10*a*, without disturbing the lattice on either side. The centre would then be in tension above and compression below the slip plane just as before. This corresponds to the Volterra 'angle' dislocation $u_\theta' = \alpha$ of figure 2, which is equivalent to the line dislocation turned through a right angle.

Figure 1 represents the condition of a crystal after slip has occurred in the (x, z) plane in the x direction, but only where x is negative, and figure 2 that of slip in the (y, z) plane in the y direction, where y is negative.

The screw dislocation of figure 4 corresponds to slip in the z direction in the (x, z) plane, again only where x is negative. Therefore if slip occurs in a limited region of a slip plane, the screw dislocation can form a boundary of the region parallel to the direction of slip, and a line (or angle) dislocation can form a boundary perpendicular to the direction of slip. A curved boundary to the region may be made up of short straight intervals of line and screw dislocations.

In three dimensions suppose a finite rectangular block in the material is moved by one atomic distance in a direction parallel to one of its sides, as in figure 11. Axes (ξ, η, ζ) are taken parallel to its edges as shown, and there are then line dislocations about ξ -axes, and line dislocations about η -axes. A curved boundary surface of a displaced region may be composed of short intervals of such dislocations.

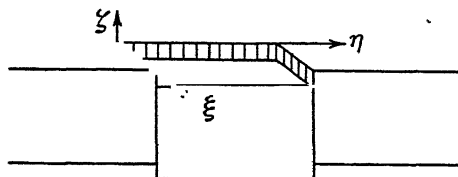


FIGURE 11

Instead of being simply displaced relative to the surrounding crystal, a certain region may suffer a uniform shear, as for example in the martensitic transformation. Figure 12 represents a rectangular block which has been sheared across (ξ, ζ) planes in the ζ direction, so that its upper surface $ABCD$ slopes down out of the (ξ, η) plane.

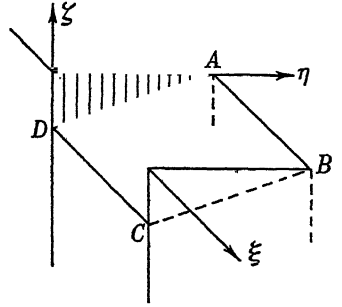


FIGURE 12

If there is only a finite number of slip planes evenly distributed between A and D , the line AD must be replaced by short steps with edges parallel to the η - and ζ -axes, and each step corresponds to a displacement as in figure 11.

As the number of slip planes increases the steps become smaller until the line AD is continuous as shown. But in the limit the linear distribution of Volterra screw dislocations may be replaced by one dislocation of the more general type $w' = \alpha r$ (figure 8). This was solved as an example in § 8 and leads to a discontinuity of the stress $\widehat{\xi\eta}$ across the (η, ζ) plane. An example of this occurred recently in some experiments by Nye.

The limit of the line dislocations about the η -axis is the linear distribution $u'_r = \alpha z$. This is similar to that shown in figure 5, but is not quite the same since there w' also exists to give a pure rotation. The pure shear $u'_r = \alpha z$, $w' = 0$, could be readily obtained by superposing the solutions corresponding to figures 5 and 8.

The limit of the line (or angle) dislocations about ξ -axes evenly distributed along AD is a single angle dislocation about the axis AB of the type $u'_\theta = \alpha r$ (figure 3).

Thus in each case the limit of a linear distribution of dislocations involving a constant discontinuity is a single dislocation with a linear discontinuity. As further examples figure 6 represents the limit of a linear distribution along the z -axis of the dislocations of figure 2, and figure 7 a linear distribution along the negative x -axis of those of figure 1.

Mathematically this is only to be expected, and has the obvious corollary that the limit of a linear distribution (along a co-ordinate line) of dislocations involving discontinuities proportional to the m th power of that co-ordinate is a single dislocation with discontinuity proportional to the $(m+1)$ th power.

10. CONCLUSION

It has been observed that plastic flow in single crystals occurs in narrow bands of high deformation separated by regions subject only to elastic stresses. These elastic stresses can now be calculated by replacing the slip bands by the appropriate dis-

locations, as indicated in this paper. It therefore becomes possible to estimate the energy stored in the material by such stresses if the amount of dislocation in the slip band is known, or, conversely, if the stresses are measured experimentally, to determine the dislocations in a slip band.

The theory given here by no means exhausts the field, but endeavours to show that quite general discontinuities in an elastic material, including those caused by plastic gliding, may be investigated by this elastic theory of dislocations.

I wish to thank Mr W. R. Dean for continual helpful suggestions and advice during the course of this work, and the Department of Scientific and Industrial Research for a grant.

REFERENCES

- Bickley, W. G. 1928 *Phil. Trans. A*, **227**, 383.
Lamb, H. 1924 *Hydrodynamics*, 5th ed. p. 47. Cambridge University Press.
Love, A. E. H. 1927a *Elasticity*, 4th ed. p. 473, Cambridge University Press.
Love, A. E. H. 1927b *Elasticity*, 4th ed. p. 327. Cambridge University Press.
Michell, J. H. 1899 *Proc. Lond. Math. Soc.* (1), **31**, 111.
Nye, J. F. To be published.
Taylor, G. I. 1934 *Proc. Roy. Soc. A*, **145**, 362.
Volterra, V. 1907 *Ann. Ec. norm. Paris*, ser. 3, **24**, 401.

The pyrolysis of methane, ethane and *n*-butane on a platinum filament

BY A. J. B. ROBERTSON, *The Davy Faraday Research Laboratory
of The Royal Institution*

(Communicated by E. K. Rideal, F.R.S.—Received 4 May 1949)

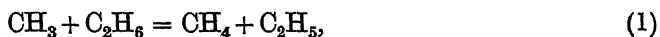
A mass spectrometer of high sensitivity and low resolving power was constructed and used to examine the nature of the reactions of methane, ethane and *n*-butane on an incandescent platinum filament at pressures of the order of 10^{-5} mm. Free radicals evolved from the filament could be directly detected by ionizing them with low-energy electrons.

The primary dissociation of methane on platinum at about 1000° C gave methyl radicals, but no methylene radicals could be detected. About 1 in 300 of the collisions of the methane on the platinum led to reaction. Further reactions of radical recombination and hydrogen transfer gave ethane, propane and butane, and part of the ethane was dehydrogenated on the platinum to ethylene. The extent of this dehydrogenation was the same as that observed with pure ethane. No free methyl or ethyl radicals could be detected in the dehydrogenation of ethane to ethylene at 950° C. Higher hydrocarbons and methane were not formed. The presence of methyl radicals from methane in the ethane did not bring about any further dehydrogenation of ethane. This reaction was concluded to be a molecular dehydrogenation. Ethyl radicals were detected from the pyrolysis of *n*-butane at 1050° C. The main products were ethane and ethylene and traces of unidentified higher hydrocarbons. No C_3 hydrocarbons and only a trace of methane were observed. Thus the butane underwent a selective fission at the central carbon-carbon bond. Consideration of bond dissociation energies in butane shows that this selective fission very probably arises from the influence of the platinum catalyst and not from a property of the butane molecule itself.

INTRODUCTION

The interpretation of the results of both homogeneous and heterogeneous pyrolysis of hydrocarbons in terms of the elementary chemical reactions involved is not yet clear. The products arising from the initial reaction are frequently less stable than the parent hydrocarbon, and may undergo extensive decomposition or reaction, thus obscuring the initial process. Even when detailed analyses are made at successive stages of decomposition, and extrapolation to zero decomposition is used to determine the products of the primary reaction (Schneider & Frohlich 1931; Neuhaus & Marek 1932) uncertainties may arise in deducing the initial mode of breakdown of the hydrocarbon. Thus, in the case of methane, the initial production of methylene radicals in the homogeneous pyrolysis has been postulated by Kassel (1932), and Belchetz & Rideal (1935) obtained methylene iodide by interaction of the products of the heterogeneous pyrolysis of methane on platinum and carbon with an iodine mirror distant about a mean free path from the origin of the products, which seemed to be direct evidence for methylene radicals. Rice & Dooley (1934) and Rice (1939), however, obtained with tellurium mirrors evidence only for methyl radicals, from both the homogeneous reaction and that on platinum. Again, Eltenton (1947), using a mass spectrometer for detecting free radicals, observed methyl, but no methylene radicals, from the decomposition of methane in a quartz vessel or on a carbon filament.

With ethane it is not certain whether the homogeneous decomposition to ethylene and hydrogen proceeds by a primary fission of the molecule to two methyl radicals, followed by the reactions



as postulated by Rice (1931), or whether a direct molecular dehydrogenation occurs, or, indeed, whether both processes of decomposition occur simultaneously (Staveley 1937). With a dehydrogenating catalyst, such as chromic oxide gel, it seems improbable that the dehydrogenation proceeds by fission to methyl radicals which escape into the gas phase, and a subsequent chain reaction. Collision of ethane molecules with a tungsten filament at high temperature, however, gives, according to Eltenton (1947), methyl radicals in the gas phase; whereas Belchetz & Rideal (1935), using a carbon filament, could find no evidence of free radical formation and concluded that the primary reaction was a molecular dehydrogenation.

The homogeneous decomposition of *n*-butane and the higher saturated hydrocarbons is generally considered to proceed largely by carbon-carbon bond fission reactions, and the extent of direct dehydrogenation to the corresponding olefins is usually small. Earlier workers represented these reactions by a series of equations for rupture of the carbon chain at any point with formation of a paraffin and olefine (Thorpe & Young 1873; Gault & Sigwalt 1927; Hague & Wheeler 1929). For *n*-butane the relative probabilities for the fission of the different bonds can be computed from the ratio $(\text{C}_2\text{H}_6 + \text{C}_2\text{H}_4)/(\text{CH}_4 + \text{C}_3\text{H}_6)$ only if chain reactions are not involved,

for, as Rice (1931) has observed, the most probable mechanism for a chain reaction is the interaction of any primarily produced free radical with butane to give either

$\text{CH}_3\text{CH}_2\text{CH}_2\text{CH}_2\text{—}$ or $\text{CH}_3\text{CH}_2\overset{|}{\text{CH}}\text{CH}_3$, which, at the temperatures used in homogeneous pyrolysis, decompose respectively, for the most part, into ethylene and an ethyl radical and propylene and a methyl radical (Frey & Hepp 1933). Interaction of these radicals with butane continues the chain, and if this is of appreciable length the products are almost entirely determined by the relative proportion of the two different butyl radicals formed and not by the point of primary rupture of the butane molecule. In fact, Frey (1934) has initiated a chain reaction in butane with methyl radicals, and observed the normal pyrolysis products.

The earlier view that the action of cracking catalysts (e.g. alumina-silica) on the higher hydrocarbons was primarily that of a source of heat (e.g. Howes 1938) has not been confirmed by detailed analytical comparison of thermal and catalytic cracking (Egloff, Morrell, Thomas & Bloch 1939; Greensfelder & Voge 1945; Haensel & Sterba 1948). The higher hydrocarbons when catalytically cracked tend to give molecules containing three or more carbon atoms. The proportion of methane and C_2 hydrocarbons is smaller in the products of catalytic cracking than in those of thermal cracking. If the mechanism at temperatures of 500 to 600° C involves acceptance of protons or hydrogen atoms by the catalyst, followed by carbon-carbon bond fission and donation of a proton or hydrogen atom to the ion or radical, in agreement with observations on deuterium exchange (Hansford 1947; Parravano, Hammel & Taylor 1948), then the catalyst must have a selective effect on carbon-carbon bond fission, since with this mechanism the distribution of products indicates the relative probabilities of fission of different bonds. The first and second bonds are relatively inert. The correlation of this selective action with the nature of the catalyst, together with the further problem of the basis of the difference between cracking and dehydrogenation catalysts, is by no means clear.

It therefore seemed of interest to examine the pyrolysis of methane, ethane and *n*-butane on a platinum filament with particular reference to the direct detection of free radicals with a mass spectrometer, under conditions such that the primary products of pyrolysis could be observed. The possibility of directly detecting free radicals with a mass spectrometer has been shown by Hipple & Stevenson (1943), and particularly by Eltenton (1942, 1947), who has also discussed the validity of the underlying assumptions of the method.

EXPERIMENTAL METHOD

The gas under examination at a pressure of 10^{-4} to 10^{-6} mm. was passed over a heated platinum spiral placed in the ionization chamber of a mass spectrometer. The products of pyrolysis and excess unchanged gas passed directly into an electron beam, where positive ions were formed. These were subjected to a mass resolution with electric and magnetic fields, and the positive-ion current for different masses was determined. When detecting free radicals, the energy of the ionizing electrons was maintained above the ionization potential of the radical, but below the critical potential (appearance potential) for the formation of the charged radical by electronic

dissociation of the parent molecule or a decomposition product. Both this method and a study of the mass spectra obtained with electrons of sufficient energy to produce extensive fragmentation of the hydrocarbons were used to investigate the formation of more stable products.

Data on ionization and appearance potentials show that in studying the formation of methyl and ethyl radicals, a mass spectrometer of high sensitivity, rather than high resolving power, is required. The pyrolysis of methane on platinum forms, as shown later, ethane, propane, butane and ethylene. Reference to data on the ionization of the methyl radical (Hipple & Stevenson 1943) and the electronic dissociation of methane (Hipple & Stevenson 1943; Smith 1937), ethane (Hipple 1938), propane (Stevenson & Hipple 1942*a*) and butane (Stevenson & Hipple 1942*b*) shows that in detecting free methyl radicals formed by the pyrolysis of methane, electrons of energy between 10 and 14 eV must be used for bringing about ionization. At these potentials the data of these authorities do not show any formation of CH_2^+ from the aliphatic hydrocarbons or ethylene, and the only contribution to the CH_4^+ ion is from methane itself, ionized at 13 eV (Honig 1948). Again, ethyl radicals must be detected below the ionization potential of ethylene (10.6 eV) to avoid interference from heavy ethylene, $^{13}\text{CH}_2\ ^{12}\text{CH}_2$. At this potential ethane is not ionized. Thus the study of free methyl and ethyl radicals by mass spectra depends on the observation of very small positive-ion peaks with neighbouring peaks which either vanish or are themselves small. The situation is different from that in work with isotopes, where a small peak must be accurately measured in relation to a much larger neighbouring peak.

The mass spectrometer

Figure 1 shows two sections of the general arrangement. The hydrocarbon flowed into tube *T* through a fine capillary tube, and passed over the heated platinum spiral *B* of geometrical surface area about 0.3 cm.². The products of pyrolysis passed through a grid covering a large hole in the electrode *C*, and thence into the electron beam *E*, emitted from the oxide-coated nickel filament *F* and controlled by the

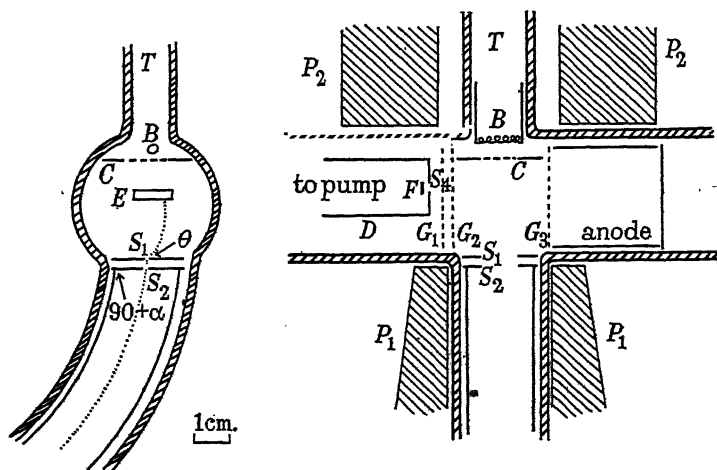


FIGURE 1. The mass spectrometer.

..... positive ion path. $\alpha = 7.5^\circ$, $\theta = 70^\circ$ approx.

electrodes D , G_1 and G_2 . To avoid interference from products formed on F , it was enclosed in the copper box D , containing a slit S_4 for the electron beam, and D was connected directly to a pumping lead. The positive ions were drawn out of the electron beam by a potential between C and an electrode containing a slit S_1 , and accelerated through S_2 into an earthed semicircular copper tube between the poles P_1 of a permanent magnet producing a field of 2100 oersteds. Ions with the proper mass/charge ratio were deflected through 180° and emerged through a slit S_3 (not shown in the figure). All the electrodes, including the whole semicircular copper tube, were enclosed in an envelope of W.1 sealing glass. All the electrodes were of copper, except the grids, which were of brass. The entire magnetic analyzer with slits S_1 , S_2 and S_3 was constructed outside the glass envelope and introduced when aligned into the glass tube. A small auxiliary electromagnet P_2 was used to straighten the lines of force in the region of the electron beam.

In examining mass spectra the initial drawing-out field between C and S_1 was kept at a constant fraction of the final accelerating field between S_1 and S_2 , so that ions of different mass would follow the same path in the source region (Jordan & Coggeshall 1942). The positive-ion current for various masses was passed through a resistance of $5 \times 10^{10} \Omega$, and the potential produced measured with an electrometer tetrode valve and galvanometers of sensitivity 450 and 4000 mm./ μA operating in the circuit of DuBridge & Brown (1933). Mercury cut-offs and cold traps were employed in the pumping and inlet leads to keep tap grease out of the mass spectrometer. The pressure in it was measured with a McLeod gauge with a compression ratio of $10^5 : 1$. The slit sizes were: $S_1 = S_2 = S_3 = 1 \times 11$ mm., $S_4 = 2 \times 9$ mm. These large slits gave adequate sensitivity, but low resolving power.

Conditions for detection of free radicals

For radicals with a molecular velocity corresponding to the temperature of the platinum filament at $1000^\circ C$, the necessary lifetime to reach the electron beam is 7 μ sec. for methyl and 10 μ sec. for ethyl. For CH_3^+ in an initial field of 65 V/cm., S_1 is reached in 0.86 μ sec., and S_2 in a further 3.9×10^{-8} sec., the ion finally emerging through S_3 into a Faraday cylinder after a further 2.3 μ sec. The $C_2H_5^+$ ion takes approximately twice as long. Thus for detection the radicals and their positive ions must be stable for several million molecular vibrations. The dissociation of the $C_2H_5^+$ ion to $C_2H_3^+$ and H_2 in this time is slight (Hipple, Fox & Condon 1946).

MATERIALS

Methane was prepared by reduction of methyl iodide with a copper-zinc couple, washed with fuming sulphuric acid and potash and dried over phosphorus pentoxide.

Ethane was prepared by hydrogenation of cylinder ethylene on a reduced nickel catalyst supported on pumice. It was washed with fuming sulphuric acid and potash, dried over phosphorus pentoxide, condensed at liquid-air temperatures and excess hydrogen pumped off.

Butane from a cylinder was dried over phosphorus pentoxide.

EXPERIMENTAL RESULTS

Pyrolysis of methane

With the platinum filament at temperatures from 920 to 1050° C (estimated with a disappearing filament pyrometer) the formation of methyl radicals was observed. No methylene radicals could be detected. In some experiments the methane was passed over the hot filament for some time and a steady state reached; in others the effect of turning the filament-heating current on and off was examined. For both methods a long preliminary heating of the filament *in vacuo* was required for quantitative observations, since it and its leads evolved gases, particularly carbon monoxide, on first heating. A temperature of 1100° C was used for this. Without the preliminary heating the gas evolution from the filament could lead to scattering of the positive-ion beam. Heating the filament in methane also led to the appearance of higher hydrocarbons, ethane, propane, butane and ethylene being detected. Acetylene was not detected. An accumulation of positive charge on the electrodes was also observed on pyrolysis of methane. All the positive-ion peaks moved slowly to a lower value of the accelerating voltage, and the electron energy between G_2 and G_3 increased, as could be seen by an apparent decrease of all the appearance potentials. This charge contamination effect was sometimes as great as 8 V. Thus it was important to check all contact potentials when observing free radicals, by reference to the ionization efficiency curve of some gas present, keeping the ion source conditions constant.

The charge contamination was only observed throughout the investigations when higher hydrocarbons were formed; it was therefore thought to be due to rather large hydrocarbon positive ions adhering, without being immediately discharged, to G_2 , G_3 and the electrode containing S_1 , and thus slightly increasing the potential of the region from which the ions were drawn out, with respect to the filament F and the earthed copper tube. When the platinum filament was cooled the peak positions drifted back to their normal values in a few minutes. The CH_3^+ current due to free radicals disappeared at once on cooling this filament.

The platinum filament gradually lost its catalytic activity with use. It could be again activated by heating at 1100° C in oxygen at a few mm. pressure for 12 hr. This poisoning seemed to be due to deposition of carbon, in agreement with the findings of Kubokawa (1937). The residual carbon dioxide in the mass spectrometer was much greater immediately after the activation than before it, in agreement with this view. The activity of the filament was somewhat greater in the first experiments on methyl radicals from methane, carried out before the wire had been given the activation treatment, than it ever became later after several activations had been carried out. Some unknown factor of the surface state of the platinum seems to be involved. The rate of catalytic combination of hydrogen and oxygen on platinum at room temperature depends on the previous treatment of the platinum (Chapman & Reynolds 1936), but with the platinum at 700 to 1900° K its previous treatment does not seem of importance, provided it is not exposed to specific poisons (Langmuir 1921). Even the exceptionally strong poisoning effect of hexamethyldisiloxane on platinum is removed by heating to 1440° K *in vacuo* (Savage 1948). Thus our

preliminary heating of the platinum at 1100° C followed by its long exposure at 1000° C to the methane should not leave sufficient foreign materials on the surface to alter the nature of the reactions, which were observed to remain qualitatively the same over a period of nearly a year, and when the platinum was heated with no preliminary outgassing. It seems unlikely, however, from electron-emission studies (Oatley 1939; Whitney 1936), that this treatment gave a truly clean platinum surface.

The steady-state partial pressures of the ethane, propane and butane formed in the mass spectrometer ionization chamber by flowing methane at a partial pressure of 2.9×10^{-5} mm. over the freshly activated platinum filament at 1020° C were determined. The positive-ion current for the $C_2H_6^+$ ion was measured to obtain the ethane pressure. A peak in the $C_4H_{10}^+$ region, identical in shape with one given by pure butane in the apparatus, was measured to obtain the butane pressure. The $C_3H_5^+$ ion was chosen for propane to avoid residual carbon dioxide. The apparatus was calibrated with ethane and butane. Table 1 gives some figures obtained.

TABLE 1. PRODUCTS OF PYROLYSIS OF METHANE

		electron energy: 50 eV			platinum filament temp. (° C)	
gas introduced	pressure (mm. mercury)	electron beam current (μA)	ion current (amp. $\times 10^{-15}$)			
			$C_4H_{10}^+$	$C_3H_5^+$		$C_2H_6^+$
methane	2.9×10^{-5}	5	15	84	300	1,020
ethane	5×10^{-5}	7.5	—	—	14,100	50
butane	3×10^{-5}	3	1,860	—	—	50

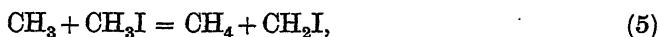
The positive-ion current from any gas is proportional to its pressure in the mass spectrometer and to the electron beam current. Taking the whole of the observed currents for $C_4H_{10}^+$ and $C_2H_6^+$, when methane is pyrolyzed, as arising from butane and ethane respectively, the calibrations give the partial pressures as 1.4×10^{-7} and 1.6×10^{-6} mm. with an error estimated as $\pm 40\%$. The contribution of the butane formed from methane to the $C_3H_5^+$ current is readily calculated from the $C_4H_{10}^+$ current to be 17.6×10^{-15} amp. (Stevenson & Hipple 1942*b*). Thus the contribution from propane is 66.4×10^{-15} amp. From the relative ionization efficiency data of Roberts & Johnsen (1948), this can be converted to a pressure of propane from the calibration figure of butane. The partial pressure of propane is then 7.3×10^{-7} mm. ($\pm 80\%$). The error here is greater since the calculation is not based on a direct calibration. Assuming the gases to be pumped away at a rate inversely proportional to the square root of their molecular weights, the molar ratio in which ethane, propane and butane were formed is 16 : 6 : 1. Ethylene was also formed, as mentioned below. The rate of flow of methane through the apparatus, at the same pressure of methane, was determined in a separate experiment. It was 9×10^{13} mol./sec. This with the analytical data enables the absolute reaction rate to be computed.

The conditions for the detection of methylene radicals are favourable, although they were not found as products of the pyrolysis. The ionization potential of the radical is about 11.9 eV (Langer & Hipple 1946), whereas the appearance potential

of CH_2^+ is 19 eV from ethylene (Kusch, Hustrulid & Tate 1937), 20 eV from methane (Koffel & Lad 1948), and 16 eV from ethane (Hipple 1938). Our findings on platinum are in agreement with those of Eltenton (1947), but do not confirm the conclusions of Belchetz & Rideal (1935). Their conclusions depend on the formation of methylene iodide on an iodine mirror. At the pressure they worked at (0.1 mm.) the reaction



(Bawn & Tipper 1947) might occur in the heated gas between the filament and iodine mirror. A further possible complicating factor is the reaction



postulated by several workers (Bawn 1939; West & Schlessinger 1938) followed by



in presence of the excess of iodine. Emschwiller (1932), indeed, has obtained methylene iodide by irradiation of methyl iodide. Thus the experimental detection of methylene iodide by Belchetz & Rideal is not inconsistent with the primary formation of methyl radicals.

Pyrolysis of ethane

The pyrolysis was examined at pressures of 1 to 7×10^{-5} mm. with temperatures of the platinum filament up to 950°C . The relative positive-ion currents due to C_2H_4^+ , C_2H_3^+ , C_2H_2^+ and C_2H^+ as compared with C_2H_6^+ were examined, using electrons of energy 50 eV. Even at the higher pressures these ions were separated sufficiently for ethylene to be determined in the presence of ethane, as may be seen from figure 2,

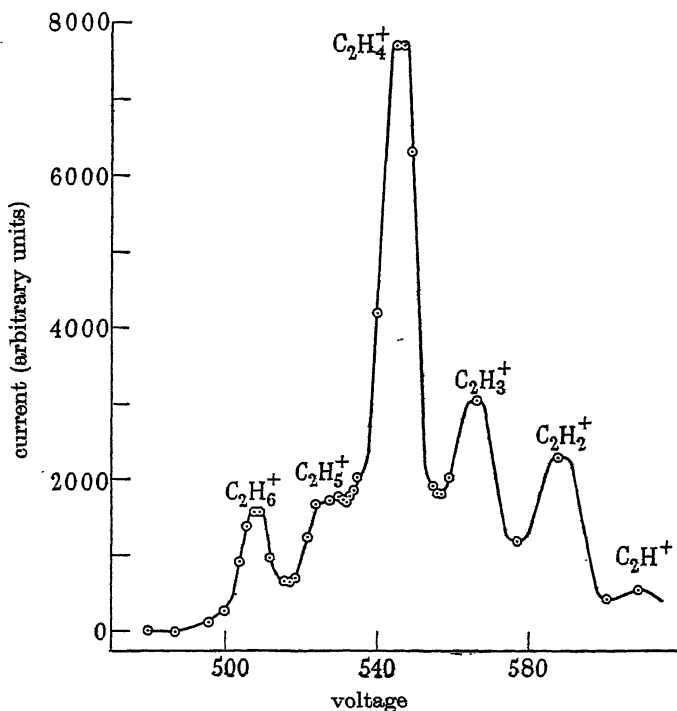


FIGURE 2. Positive ion curve given by ethane.

showing a mass spectrum of ethane at 5×10^{-5} mm. pressure. The separation of $C_2H_5^+$ from $C_2H_4^+$ was very poor, and no quantitative conclusions were drawn from measurement of the $C_2H_5^+$ current. Comparison of data obtained with the platinum filament cold and at 950°C showed that ethane present at these pressures was dehydrogenated on the hot filament to ethylene to the extent of about 20 %. It was not necessary to activate the filament to bring this about. The formation of methane and hydrocarbons of greater mass than ethane was not observed, although the reaction of 0.2 % of the ethane to give these products should have been observable. No acetylene was observed, but the sensitivity here was less. The charge contamination effect observed on pyrolyzing methane was not observed with ethane. A direct search for methyl and ethyl radicals with low-energy electrons failed to reveal any with the platinum filament at 950°C . Thus the observed effect of the platinum filament on ethane was to bring about a molecular rearrangement to ethylene and hydrogen. Any dissociation to methyl or ethyl radicals, which could appear in the gas phase, if it occurred at all, must have been to a very limited extent.

Comparison of ethylene formation from ethane and methane

Using the same electrical conditions and filament temperature, a comparison was made of the relative formation of ethylene from ethane and methane. With the filament freshly activated and at temperatures up to 950°C the ratio of the ion currents due to $C_2H_6^+$ and $C_2H_4^+$ was observed, with 50 eV electrons. These ions were used to avoid interference by residual carbon monoxide. Within the limits of error, this ratio was the same when ethane and methane were pyrolyzed. Thus the bulk of the ethylene formed from methane could be accounted for as a secondary product formed by dehydrogenation of ethane. In these experiments no calibration was required, and thus errors were minimized. It was estimated that not more than 20 % of the ethylene formed from methane was formed in some other way.

Pyrolysis of n-butane

Here the possible primary reactions are dehydrogenation and fission of the central or end carbon-carbon bonds. At a pressure of 3×10^{-5} mm. and with a filament temperature of 1050°C about 16 % of the butane decomposed, according to the diminution of the positive-ion current in the $C_4H_{10}^+$ and $C_3H_7^+$ regions, on heating the previously outgassed filament. Ethane and ethylene were the only two major products observed in the mass range covering C_4 to C_1 hydrocarbons. Measurement of the $C_2H_6^+$ and $C_2H_4^+$ currents, and correction for the contributions of butane and ethane to this latter, gave the molar ratio in which ethane and ethylene were present as 1 : 1.5 ($\pm 50\%$). Thus when allowance is made for the secondary dehydrogenation of the ethane, these products were formed in roughly equimolar amounts. The sensitivity for C_3 products was not very great, since the butane contributed greatly to the various C_3 ion currents. No clear evidence for the formation of propane or propylene could be obtained. The formation of a trace of methane was observed by study of mass 16. The conclusion reached was that decomposition to ethane and ethylene was at least five times more probable than decomposition to methane and propylene, and the ratio was probably greater.

The extent of dehydrogenation of the butane to butene or butadiene was investigated by examining the shape of the positive-ion curve in the C_4 region with the platinum filament hot and cold. The ordinary method of analysis could not be used because of the limited resolving power of the apparatus. Figure 3 shows a curve obtained with the filament at the temperature of the mass spectrometer (50°C), and figure 4 a curve with the filament at 1050°C . The two curves are practically identical in shape. The position of the maximum is displaced by 5 V with the filament at 1050°C by the charge contamination effect, which was again observed on pyrolysis of butane. Formation of butenes would be revealed by a bulge on the side of the curve at a voltage corresponding to mass 56. Such a bulge is visible (because of the electronic dissociation), but it is not increased by pyrolyzing the butane. To make a more sensitive test for butenes, electrons of very low energy were used so that the $C_4H_{10}^+$ peak was just observable. Since the ionization potential of the various butenes is 1 eV or more below that of butane, any appreciable dehydrogenation of the butane on the platinum and evaporation of butenes should then have produced a considerable change in the shape of the positive-ion curve. Neither the butenes nor butadiene were detected. Thus the decomposition involved carbon-carbon bond fission, as is also the case for the homogeneous pyrolysis. This bond fission on platinum seems to be rather selective, and if carbon-hydrogen bonds are broken and butene molecules formed, these do not evaporate before undergoing further reaction.

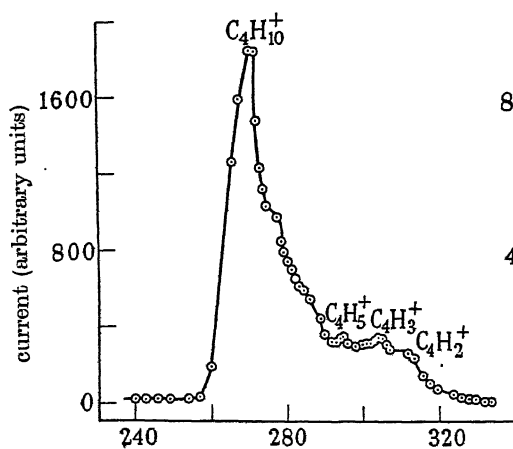


FIGURE 3. Positive ion curve from butane.
50 eV electrons.

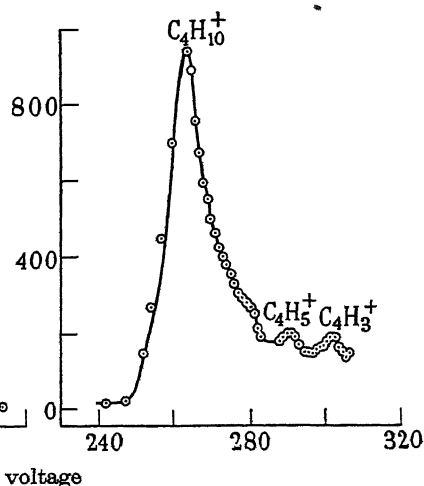


FIGURE 4. Positive-ion curve from
pyrolyzed butane.

Higher hydrocarbons were searched for in the mass range 60 to 100. The resolving power in this region was small, but two distinct peaks appeared on pyrolyzing butane. Figure 5 shows an ion curve obtained with 35 eV electrons (to diminish the Hg^{++} peak) with butane at 1×10^{-5} mm. pressure passing over the platinum filament at 1050°C . The currents due to higher hydrocarbons are very small compared with that of the $C_4H_{10}^+$ ion. There are two distinct maxima in the current curve, corresponding within 3 mass units, to singly charged ions of mass 84–86 and 70–71, the

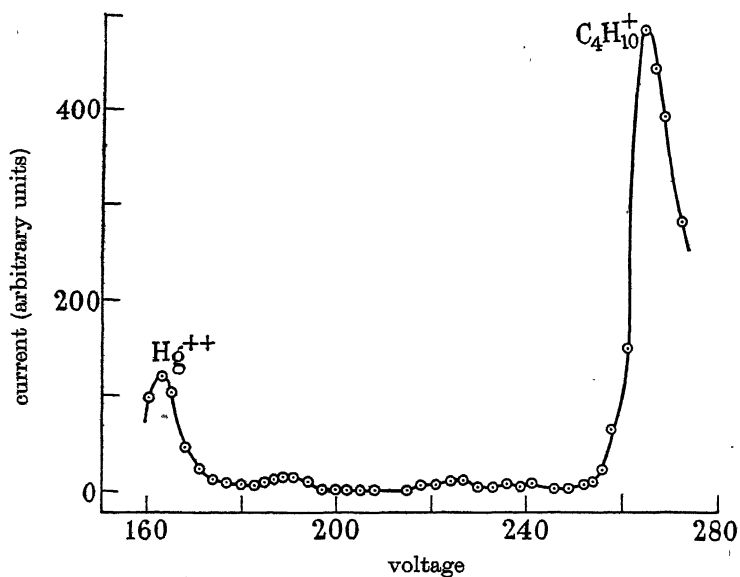


FIGURE 5. Positive-ion curve showing formation of higher hydrocarbons on pyrolysis of butane.

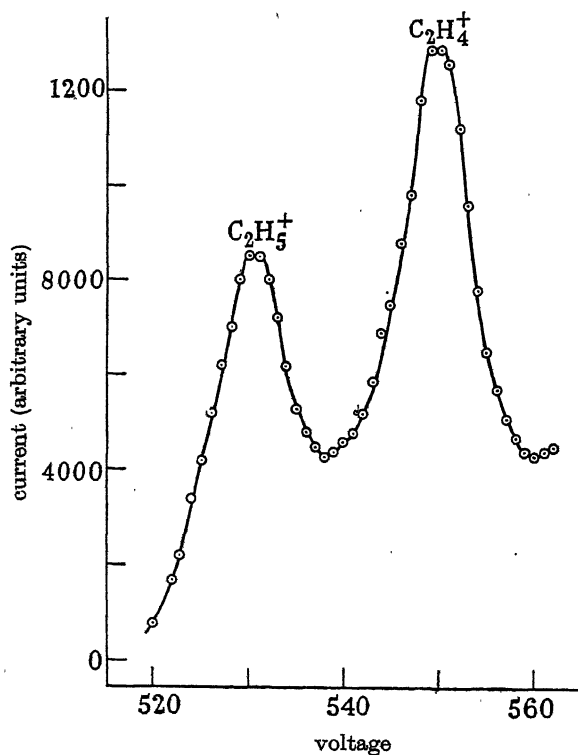


FIGURE 6. Separation of $C_2H_4^+$ and $C_2H_5^+$ in the mass spectrum of pyrolyzed butane, using 50 eV electrons.

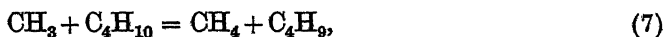
ratio of currents being 1:0.8. Peaks at these masses are prominent in the mass spectra of several of the C_6H_{14} and C_8H_{18} hydrocarbons. With the limited resolving power an identification of any C_6 or C_8 hydrocarbons present could not be made, but it is clear that at least one higher hydrocarbon must be formed.

A direct search for free ethyl radicals was made. The $C_2H_4^+$ and $C_2H_5^+$ currents were sufficiently separated in the mass spectrum of pyrolyzed butane to enable them to be compared, as may be seen from figure 6. The currents were compared at diminishing electron energies until the $C_2H_4^+$ current vanished. On increasing the butane pressure in the ionization chamber to almost 10^{-4} mm., a small current due to $C_2H_5^+$ was detected at a potential nearly 0.5 eV lower than the ionization potential of ethylene. A current of nearly 1×10^{-15} amp., which must have been due to ethyl radicals, was obtained with a filament temperature of 1050° C. Therefore the heterogeneous pyrolysis formed free ethyl radicals which could escape into the gas phase.

DISCUSSION

The very low pressure of the reactant and the short distance between the incandescent filament bringing about the reaction and the electron beam and its associated electrodes are favourable for the study of the nature of the primary step in the pyrolysis of the hydrocarbons, but they are not favourable for the exact study of the various secondary reactions which may be brought about by free radicals liberated from the filament. In general, any such radicals will collide with a surface in the mass spectrometer more frequently than with a molecule in the gas phase, since their mean free path is considerably greater than the dimensions of the apparatus. The secondary reactions, such as radical recombination, probably occur very largely as surface reactions (Paneth & Lautsch 1931). The conditions are somewhat arbitrary for any detailed study of these reactions, since many different kinds of surface are present in the spectrometer and their state is uncertain. Thus the results of studies of the homogeneous gas-phase reactions of free radicals are not directly comparable with those given here.

Direct detection of free radicals shows that methane on an active platinum filament at about 1000° C is dissociated to methyl radicals and no methylene radicals can be detected. The formation of propane and butane from methane shows that some ethyl radicals are probably present in the system, and radical recombination of methyl and ethyl gives ethane, propane and butane. A probable mechanism for the formation of ethyl radicals is reaction (1). The ethane itself does not give observable ethyl radicals on reacting on the platinum, but forms ethylene. It is conceivable that a further source of ethyl radicals in the methane system is the reaction of this ethylene on the platinum wire with hydrogen atoms from methane, followed by evaporation of an ethyl radical. Further reactions of the type of (1), e.g.



probably account for the charge contamination of the electrodes which is observed when methane is pyrolyzed. This effect was not observed when butane was present

in the mass spectrometer with the platinum wire cold, and still larger hydrocarbons are probably formed from methane.

Bimolecular recombination reactions such as (8) in the gas phase have low-activation energies and very small steric factors (Bawn 1935). Hydrogen transfer reactions such as (1), (3) and (7) also have low activation energies (Rice & Teller 1938; Smith & Taylor 1939). Thus our experimental conditions favour such reactions as compared with radical decomposition reactions such as (2) requiring appreciable activation energies. Some of the methyl radicals formed from methane probably react with ethylene



the propyl radical giving propane by reaction with methane, and combining with methyl radicals to form butane. Eltenton (1947), however, has shown that methyl radicals disappear more rapidly in ethane than in ethylene. Since more ethane than ethylene is present when methane is pyrolyzed, (1) should be more frequent than (9). Thus the yield of propane is considerably greater than that of butane. The absence of butane when ethane is pyrolyzed shows that it is not formed by combination of ethane and ethylene on the hot platinum.

The absolute rate of the dissociation of methane can be calculated from the experimental results. The flow of gas away from the mass spectrometer ionization chamber is molecular streaming through a tube; therefore, a gas of molecular weight M at a partial pressure p is pumped away at a rate proportional to $p/\sqrt{M}$. By experiment, methane at a pressure of 2.9×10^{-5} mm. is pumped away at a rate of 9×10^{13} mol./sec. Therefore ethane, at its observed pressure, when methane is pyrolyzed, of 1.6×10^{-6} mm. is pumped away at a rate of 3.6×10^{12} mol./sec. Since a steady state is reached, the rate of transformation of methane to ethane is 7.2×10^{12} mol./sec. With 20 % of the ethane formed dehydrogenated to ethylene, the rate of transformation of methane to ethane plus ethylene is 9×10^{12} mol./sec. By a similar calculation for propane and butane, the total rate of removal of methane is 1.4×10^{13} mol./sec. The number of collisions of methane with the platinum is 4.3×10^{15} mol./sec. Thus 1 collision in 310 brings about reaction, assuming the true area of the platinum to equal its geometrical area. This figure is subject to uncertainty by a factor of 2 or 3. The absolute rate of reaction, however, is clearly very great. The mechanism cannot involve thermal activation of the methane molecules on the platinum, followed by a monomolecular decomposition in the gas phase. In such a case the molecules would have to decompose in about $10 \mu\text{sec.}$, before they were deactivated by collision with a cold surface. The fraction able to decompose in this time should be of the order of $10^{-5} \times 10^{13} \exp(-102 \text{ kcal.}/RT)$, since the C-H bond dissociation energy in methane is 102 kcal./g. mol. at 20°C (Kistiakowsky & Van Artsdalen 1944). At 1000°C this fraction is $10^{-9.5}$. The platinum clearly exerts a most powerful catalytic effect on the decomposition of the methane. Some experiments of Beeck (1935, 1948) also reveal a great catalytic effect of platinum on some hydrocarbon reactions at low pressure, and he states that the presence of water vapour facilitates this effect. In all our experiments, water vapour could be detected in the mass spectrometer, and this may be significant in considering our high absolute reaction rate.

The ethylene formed when methane is pyrolyzed can, within the limits of experimental error, be entirely ascribed to the secondary reaction of the dehydrogenation of ethane formed by the combination of the primarily generated methyl radicals, if it is assumed that the efficiency of the ethane dehydrogenation is the same in the presence of excess methane as in the ethane-ethylene-hydrogen system. The efficiency may be slightly less, because when methane is present some hydrogenation of ethylene by atomic hydrogen could occur. This effect, however, is probably counter-balanced by small contributions to the formation of ethylene in the methane system from decomposition of propane and butane. The contribution to the formation of ethylene in the methane system by other mechanisms than that of the dehydrogenation of ethane cannot be great. The extent of dehydrogenation of the ethane is much less than that corresponding to the ethane-ethylene-hydrogen equilibrium, which at these very low pressures is almost wholly in favour of ethylene and hydrogen. Thus it must be determined by the reaction rate of ethylene formation. The contribution to this from decomposition of propane and butane and from disproportionation of ethyl radicals cannot be great. In fact, the experimental data show that the dehydrogenation of ethane at 950° C occurs to a greater extent than the decomposition of *n*-butane at 1050° C. No quantitative estimate of the relative importance of combination and disproportionation of ethyl radicals can be made from the data, since the yield of butane is smaller than that of ethylene, and butane is probably formed both from propyl and methyl radicals and from ethyl radicals. The disproportionation of methyl radicals to methane and methylene radicals, followed by dimerization of methylene to ethylene, cannot be an important source of ethylene. This is also true for the homogeneous gas phase reactions of methyl radicals (Bawn & Tipper 1947). Again, the methyl radicals formed from methane do not react to any significant extent according to

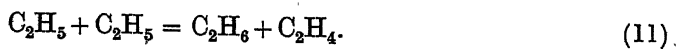


The decomposition of ethyl radicals according to (2) must occur to only a slight extent compared with the radical association reactions of ethyl with itself and methyl.

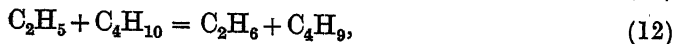
When methane is pyrolyzed, methyl radicals are present, and the existence of ethyl radicals may be inferred, but pyrolysis of ethane alone gives insufficient of either radical for detection. The ethane formed from methane also undergoes pyrolysis on the filament, but in this case methyl radicals are present from methane. Since the ethane/ethylene ratio is the same when pure ethane and methane are pyrolyzed, the introduction of methyl radicals into ethane at the very low pressures used does not bring about much, if any, further dehydrogenation of the ethane to ethylene. Thus, since the platinum filament brings about dehydrogenation of ethane, the primary reaction of the ethane on the platinum cannot be a fission to methyl radicals which evaporate into the gas phase. Thus in this case the scheme of Rice, whereby the dehydrogenation proceeds by a primary rupture to two methyl radicals, is by no means supported. Further, of the secondary reactions (1), (2) and (3) of Rice, the experimental evidence is against the occurrence of (2) to any significant extent when ethyl is present, and since methane is not formed appreciably, (1) does

not occur when ethane is pyrolyzed. All the results with methane lead to the conclusion that a primary scission of ethane to methyl radicals would be followed by (1) and association reactions giving propane and butane, and higher hydrocarbons by reactions such as (7) and (8). This was not observed. Again, no charge accumulates on the electrodes when ethane is pyrolyzed, which seems to indicate that very large hydrocarbons are not formed. Thus one may conclude that the heterogeneous reaction of the formation of ethylene from ethane on platinum at about 950° C proceeds very largely as a direct molecular dehydrogenation on the catalyst, and free radicals are not evolved from the catalyst. This conclusion is in agreement with that reached by Belchetz & Rideal (1935) from their investigation of the pyrolysis on carbon at 1450 to 1800° C, and with the reverse molecular addition of hydrogen to ethylene in catalytic hydrogenation postulated by Twigg & Rideal (1939) and Farkas & Farkas (1937).

Steacie (1946), reviewing the homogeneous pyrolysis of methane and ethane, concludes that the primary step for methane is the formation of methyl radicals, whereas for ethane molecular dehydrogenation occurs more frequently than the formation of two methyl radicals. Thus the catalytic reactions on platinum show no specific features. With *n*-butane on platinum, however, the main products are ethane and ethylene, and ethyl radicals can be detected. Only a trace of methane is observed, and methyl radicals are therefore probably not formed to a significant extent. These results clearly point to the fission of the butane molecule at the central C-C bond, to give two ethyl radicals, or ethane and ethylene. The ethyl radicals may react with themselves to give butane again, or by disproportionation



The reactions



with reaction (8), could give hexane and octane, and similar reactions probably give higher hydrocarbons, thus accounting for the charging of the electrodes. *n*-Butane reacting by (12) should form mainly $\text{CH}_3\text{CH}_2\text{CHCH}_3$, which decomposes thermally to propylene and a methyl radical. These are not detected, and thus this decomposition cannot be a significant reaction in our experiments. Thus again with butane the experiments show the predominance of radical association reactions and hydrogen transfer reactions over radical decomposition reactions at the very low pressures used.

It is not possible to say whether a dissociation to ethyl radicals which then evaporate is the only primary step in the butane decomposition without knowledge of the relative probabilities of the occurrence of (11) and (12). (11) would have to be rather more probable than (12) to account for the approximately equimolar formation of ethane and ethylene, and the small yield of hydrocarbons higher than butane. From the experiments with methane and ethane this does not seem very likely, and some ethylene is probably formed on the filament when butane is pyrolyzed.

It is not entirely certain that a specific catalytic effect must be invoked for the butane pyrolysis. The fact that the catalytic pyrolysis results largely in C-C bond

fission rather than dehydrogenation shows an analogy with the homogeneous pyrolysis (Steacie & Puddington 1938; Echols & Pease 1939). Hinshelwood (1948) calculates from the analytical results of Steacie & Folkins (1940) on the nitric oxide inhibited decomposition of *n*-butane that the probabilities of fission according to (14) and (15) are about equal:



Steacie & Folkins, however, consider that the nitric oxide merely diminishes the chain length in the reaction without actually suppressing the chain. In this case the primary rupture cannot be deduced with any certainty from the ratio

$$(\text{CH}_4 + \text{C}_3\text{H}_6)/(\text{C}_2\text{H}_6 + \text{C}_2\text{H}_4).$$

Thus from the experimental viewpoint it is not clear how the primary rupture of a butane molecule in the gas phase occurs. If the radical recombinations of (14) and (15) are assumed to have no activation energy, the bond dissociation energies give the activation energies of the two decompositions. If in the homogeneous decomposition of butane at 1000° C the ratio of fission according to (14) and (15) is $\exp(-\Delta E/RT)$, where ΔE is the activation energy difference, ΔE requires to be 6 kcal. for the central bond fission to be 10 times more probable than the end bond fission. This seems, in the light of the bond energy considerations of Roberts & Skinner (1949), an unlikely value for ΔE , and thus the observed selective fission probably arises from some influence exerted by the catalyst and not from a property of the butane molecule itself.

One possibility is that the butane may lose one hydrogen atom from each methyl group on adsorption, and the reaction



will then clearly be favoured over other bond fission reactions. Bawn & Milsted (1939) estimate from energy surfaces an activation energy of only 15 kcal. for the homogeneous reaction (16). On platinum, adsorption of ethylene might make this even lower. The ethylene could evaporate as such, or be hydrogenated by the neighbouring atomic hydrogen on the surface to give ethyl and ethane, which also evaporate. This scheme, although speculative, is in conformity with the views on catalytic cracking advanced by Parravano *et al.* (1948). Platinum has a face-centred cubic lattice, the length of the side of the elementary cube being 3.93 Å (Hull 1921). For *n*-butane fully extended in a plane, the distance between the two end carbon atoms is 3.9 Å. A possible configuration of the *n*-butane skeleton in relation to a platinum surface is shown in figure 7. It is clear from this that a small rotation about the middle C-C bond would permit of interaction between the two platinum atoms *A* and *C* and C-H bonds in each methyl group of the butane. Even if this interaction does not lead to breakage of a C-H bond, it seems possible that interaction of the two CH₂ groups with the platinum atoms *B* and *D* could lead to sufficient strain on the central C-C bond to bring about a direct fission of the molecule to ethyl radicals.

The author has pleasure in thanking Mr W. N. Dear for practical solution of the glass-blowing problems in constructing the mass spectrometer, Mr D. Hadfield and Messrs William Jessop and Sons for designing and constructing the permanent magnet, the Managers of the Royal Institution for the facilities afforded at the Davy Faraday Laboratory, and Professor E. K. Rideal for his continued interest and encouragement.

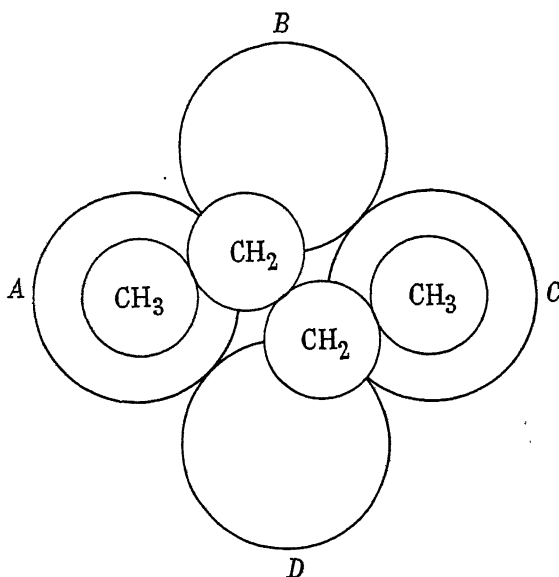


FIGURE 7. A possible configuration of the butane carbon skeleton on platinum.

REFERENCES

- Bawn, C. E. H. 1935 *Trans. Faraday Soc.* **31**, 1536.
 Bawn, C. E. H. 1939 *Trans. Faraday Soc.* **35**, 898.
 Bawn, C. E. H. & Milsted, J. 1939 *Trans. Faraday Soc.* **35**, 889.
 Bawn, C. E. H. & Tipper, C. F. H. 1947 *Disc. Faraday Soc.* **2**, 104.
 Beeck, O. 1935 *Nature*, **136**, 1028.
 Beeck, O. 1948 *Rev. Mod. Phys.* **20**, 127.
 Belchetz, L. & Rideal, E. K. 1935 *J. Amer. Chem. Soc.* **57**, 1168.
 Chapman, D. L. & Reynolds, P. W. 1936 *Proc. Roy. Soc. A*, **156**, 284.
 DuBridge, L. A. & Brown, H. 1933 *Rev. Sci. Instrum.* **4**, 532.
 Echols, L. S. & Pease, R. N. 1939 *J. Amer. Chem. Soc.* **61**, 208.
 Egloff, G., Morrell, J. C., Thomas, C. L. & Bloch, H. S. 1939 *J. Amer. Chem. Soc.* **61**, 3571.
 Eltenton, G. C. 1942 *J. Chem. Phys.* **10**, 403.
 Eltenton, G. C. 1947 *J. Chem. Phys.* **15**, 455.
 Emschwiller, G. 1932 *Ann. Chim. (Phys.)*, **17**, 413.
 Farkas, A. & Farkas, L. 1937 *Trans. Faraday Soc.* **33**, 837.
 Frey, F. E. 1934 *Industr. Engng Chem.* **26**, 198.
 Frey, F. E. & Hepp, H. J. 1933 *J. Amer. Chem. Soc.* **55**, 3357.
 Gault, H. & Sigwalt, R. 1927 *Ann. Off. Combust. liq.* **2**, 309, 543.
 Greensfelder, B. S. & Voge, H. H. 1945 *Industr. Engng Chem.* **37**, 514.
 Haensel, V. & Sterba, M. J. 1948 *Industr. Engng Chem.* **40**, 1660.
 Hague, E. N. & Wheeler, R. V. 1929 *J. Chem. Soc.* p. 378.
 Hansford, R. C. 1947 *Industr. Engng Chem.* **39**, 849.
 Hinshelwood, C. N. 1948 *J. Chem. Soc.* p. 531.
 Hipple, J. A. 1938 *Phys. Rev.* **53**, 530.

- Hipple, J. A., Fox, R. E. & Condon, E. U. 1946 *Phys. Rev.* **69**, 347.
Hipple, J. A. & Stevenson, D. P. 1943 *Phys. Rev.* **63**, 121.
Honig, R. E. 1948 *J. Chem. Phys.* **16**, 105.
Howes, D. A. 1938 *The science of petroleum*, p. 2119. Oxford University Press.
Hull, A. W. 1921 *Phys. Rev.* **17**, 571.
Jordan, E. B. & Coggeshall, N. D. 1942 *J. Appl. Phys.* **13**, 539.
Kassel, L. S. 1932 *J. Amer. Chem. Soc.* **54**, 3949.
Kistiakowsky, G. B. & Van Artsdalen, E. R. 1944 *J. Chem. Phys.* **12**, 469.
Koffel, M. B. & Lad, R. A. 1948 *J. Chem. Phys.* **16**, 420.
Kubokawa, M. 1937 *Rev. Phys. Chem. Japan*, **11**, 82.
Kusch, P., Hustrulid, A. & Tate, J. T. 1937 *Phys. Rev.* **52**, 843.
Langer, A. & Hipple, J. A. 1946 *Phys. Rev.* **69**, 691.
Langmuir, I. 1921 *Trans. Faraday Soc.* **17**, 621.
Neuhaus, M. & Marek, L. F. 1932 *Industr. Engng Chem.* **24**, 400.
Oatley, C. W. 1939 *Proc. Phys. Soc. Lond.* **51**, 318.
Paneth, F. & Lautsch, W. 1931 *Ber. dtsh. chem. Ges.* **64** B, 2708.
Parravano, G., Hammel, E. F. & Taylor, H. S. 1948 *J. Amer. Chem. Soc.* **70**, 2269.
Rice, F. O. 1931 *J. Amer. Chem. Soc.* **53**, 1959.
Rice, F. O. 1939 *J. Amer. Chem. Soc.* **61**, 213.
Rice, F. O. & Dooley, M. D. 1934 *J. Amer. Chem. Soc.* **56**, 2747.
Rice, F. O. & Teller, E. 1938 *J. Chem. Phys.* **6**, 489.
Roberts, J. S. & Skinner, H. A. 1949 *Trans. Faraday Soc.* **45**, 339.
Roberts, R. H. & Johnsen, S. E. J. 1948 *Analyt. Chem.* **20**, 690, 1225.
Savage, R. H. 1948 *J. Chem. Phys.* **16**, 237.
Schneider, V. & Frohlich, P. K. 1931 *Industr. Engng Chem.* **23**, 1405.
Smith, L. G. 1937 *Phys. Rev.* **51**, 263.
Smith, J. O. & Taylor, H. S. 1939 *J. Chem. Phys.* **7**, 390.
Staveley, L. A. K. 1937 *Proc. Roy. Soc. A*, **162**, 557.
Steacie, E. W. R. 1946 *Atomic and free radical reactions*. New York: Rheinold Corp.
Steacie, E. W. R. & Folkins, H. O. 1940 *Canad. J. Res. B*, **18**, 1.
Steacie, E. W. R. & Puddington, I. E. 1938 *Canad. J. Res. B*, **16**, 176.
Stevenson, D. P. & Hipple, J. A. 1942a *J. Amer. Chem. Soc.* **64**, 2769.
Stevenson, D. P. & Hipple, J. A. 1942b *J. Amer. Chem. Soc.* **64**, 1588.
Thorpe, T. E. & Young, J. 1873 *Proc. Roy. Soc.* **21**, 184.
Twigg, G. H. & Rideal, E. K. 1939 *Proc. Roy. Soc. A*, **171**, 55.
West, W. & Schlessinger, L. 1938 *J. Amer. Chem. Soc.* **60**, 961.
Whitney, L. V. 1936 *Phys. Rev.* **50**, 1154.

Electromagnetic induction in a rotating sphere

By E. C. BULLARD, F.R.S., *University of Toronto*

(Received 10 May 1949)

An account is given of the induction of electric currents in a rotating conducting sphere surrounded by a stationary conducting shell. Both induction by a constant field and the freely decaying modes are treated. The associated mechanical couples are discussed. The results of most interest are those concerning the toroidal magnetic fields. These fields are entirely confined to the interior of the conductor and may play an important part in terrestrial magnetism. Induction in an oscillating sphere is briefly discussed.

I. INTRODUCTION

Electromagnetic induction in a stationary sphere has been discussed by several authors. The results are summarized by Chapman & Bartels (1940, vol. 2, pp. 732–749). A rotating sphere surrounded by an insulator has been treated by Hertz, Thomson (1893) and others. More recently, Elsasser (1946*a, b*, 1947) has considered the problem of induction for any type of motion possible in a liquid sphere. He is mainly concerned with the symmetry shown by the solutions, and does not give a detailed discussion of particular motions.

The results produced by induction in moving media are so varied and interesting and have such important applications, that a detailed discussion of simple cases by exact methods is desirable. This paper treats induction in a rotating sphere. The work differs from that of the earlier workers by supposing the sphere to be surrounded by a conducting shell. This system has solutions that do not occur when the rotating sphere is surrounded by an insulator. The primary reason for making these calculations was to apply the results to problems in terrestrial magnetism (Bullard (1948), which is referred to hereafter as I, and Bullard (1949), which is referred to as II). The choice of numerical examples has been largely dictated by this interest, but the analytical discussion covers a wider field.

The most important practical results are the expressions (25) for the toroidal field and (29) for the mechanical couple produced by induction with a field parallel to the axis of rotation.

2. THE EQUATIONS FOR A SPHERE

The system for which a solution is required consists of a sphere of radius b and uniform electrical conductivity κ , rotating with angular velocity ω about an axis through its centre. The sphere is surrounded by a concentric spherical shell of the same conductivity and with an outer radius a . The sphere and the shell are supposed in perfect electrical contact over their common surface. As particular cases, the shell may be absent ($b = a$), or may fill the whole of the space not occupied by the sphere ($a \rightarrow \infty$).

The solutions of the electromagnetic equations for such a system are of two kinds. There are oscillations having a period comparable with the time taken for light to

cross the sphere, and there are fields that are stationary or decay exponentially over much longer times. For the problems considered in II the periods of the oscillations are of the order of 0.05 sec., whilst the times of relaxation of the decaying fields are of the order of 10^4 yr. The oscillations are not essentially affected by rotations that are slow compared with the velocity of light. They have been widely discussed in radio and optical applications, and will not be considered further here. The stationary and slowly varying fields are profoundly affected by quite slow rotations and are the subject of this paper. For them the displacement current plays no essential part and can be omitted from the equations. With this simplification Maxwell's equations for a moving medium become (Tolman 1934, pp. 105–109)

$$\text{curl } H = 4\pi I = 4\pi\kappa(E + v \wedge H), \quad (1)$$

$$\text{curl } E = -\dot{H}, \quad (2)$$

$$\text{div } H = 0, \quad (3)$$

$$\text{div } E = 4\pi\rho c^2/\epsilon, \quad (4)$$

where H is the magnetic field, E the electric field, I the current density, κ the electrical conductivity and ρ the charge density, all in electromagnetic units. v is the velocity of the medium at any point, ϵ its dielectric constant, and c is the velocity of light. In these equations relativistic terms of order v^2/c^2 and terms in ρv , corresponding to the convection of charge, have been omitted. It has also been assumed that the material has unit permeability; this assumption simplifies the subsequent calculations somewhat, and is allowable in the applications of II.

Eliminating E from (1) and (2) gives

$$-\text{curl}^2 H = 4\pi\kappa\{\dot{H} - \text{curl}(v \wedge H)\}. \quad (5)$$

If all distances are measured in terms of the radius of the rotating sphere and all times in units of $4\pi\kappa b^2$, this equation becomes

$$-\text{curl}^2 H = \dot{H} - \text{curl}(v_1 \wedge H),$$

where $v_1 = 4\pi\kappa b v$ is the velocity in the same units. Thus two similar systems differing only in size will have equal fields if the motions are proportioned so that $\kappa b v$ is the same at corresponding points in the two systems. The ratio of the fields at any two points in a system with a prescribed type of motion and prescribed boundary conditions at infinity will be a function of $\kappa b v_0$ only, where v_0 is any velocity characteristic of the motion, such as the mean velocity, or the maximum velocity. These theorems are quite general and apply to any motion in a body of any shape. If the body is not spherical, b can be taken as any dimension defining its size.

If the motion is one of rotation with angular velocity ω ,

$$v = \omega \wedge r;$$

substituting this in (5) and using spherical polar co-ordinates (r, θ, ϕ) , where ω lies along the axis of the co-ordinates, gives

$$-\text{curl}^2 H = 4\pi\kappa(\omega \partial H / \partial \phi + \dot{H}). \quad (6)$$

When an H has been found satisfying (3) and (6), E can be obtained from (1) or (2) and ρ from (4).

The equation (6) is not changed if the system of co-ordinates rotates uniformly about its axis. If, therefore, the shell is considered as stationary this only implies that ω in (6) is measured relative to the shell. In the applications of II the co-ordinate system rotates with the outer part of the earth's core and ω represents the difference in the angular velocities of the interior and exterior parts of the core; it is not the whole angular velocity of rotation of the earth.

3. SOLUTIONS OF (6)

$$\text{Let} \quad H = H_1 e^{im\phi} e^{-\sigma t}, \quad (7)$$

where H_1 is a function of r and θ only, m is an integer and σ a constant which may be complex or real; (6) then becomes

$$-\text{curl}^2 H = 4\pi\kappa(i\omega m - \sigma)H$$

$$\text{or} \quad (-\text{curl}^2 + k^2)H = 0, \quad (8)$$

$$\text{where} \quad k^2 = 4\pi\kappa(\sigma - i\omega m). \quad (9)$$

k may be regarded as an adjustable constant which can be chosen to make σ zero in a problem of induction by a given constant field. In a problem of a freely decaying field k may be chosen to fit the boundary conditions; σ will then be fixed by (9).

It has been shown by Lamb (1881) that the solutions of (8) that also satisfy (3) are

$$H_T = -r \wedge (\nabla\psi) \quad (10)$$

$$\text{and} \quad H_S = k^{-1} \text{curl } H_T, \quad (11)$$

$$\text{where} \quad \psi = (kr)^{-\frac{1}{2}} J_{n+\frac{1}{2}}(kr) P_n^m(\theta) e^{im\phi} e^{-\sigma t}$$

$$\text{or} \quad = (kr)^{-\frac{1}{2}} Y_{n+\frac{1}{2}}(kr) P_n^m(\theta) e^{im\phi} e^{-\sigma t}.$$

From (8) and (11) H_T and H_S are also related by

$$H_T = k^{-1} \text{curl } H_S. \quad (12)$$

In Elsasser's notation† (10) is the toroidal field T_n^m and (11) is the poloidal field S_n^m . These solutions are in general complex; real ones (Elsasser's T_n^{mc} , T_n^{ms} , S_n^{mc} and S_n^{ms}) can be obtained by adding a pair of conjugate solutions. For a given n and m there are in all eight real solutions, four of which are toroidal and four poloidal. Each set of four has two with a singularity at the origin and two without. For real k each of these pairs has one solution in $\cos m\phi$ and a term in $\sin m\phi$, but if k is complex such a division is only possible if the zero for ϕ is correctly chosen.

Equation (11) shows that the current corresponding to a magnetic field H_T is $kH_S/4\pi$; that is, a toroidal field T_n^m is accompanied by a poloidal current. (12) shows that the converse is also true.

It is desirable to have a notation for the field and current vectors that will not only distinguish between (10) and (11), but will also indicate the components of the various quantities. We use ${}_r H_{Tnm}$ to indicate the r component of the magnetic field

† To obtain symmetry between (11) and (12) and a non-dimensional form for ψ , the definitions of S_n^m and T_n^m have been slightly changed.

given by the solution (10) containing $J_{n+\frac{1}{2}}(kr)$ and similarly for the θ and ϕ components. The solution (11) is written as ${}_rH_{Snm}$. The solutions with $Y_{n\pm\frac{1}{2}}(kr)$ in place of $J_{n\pm\frac{1}{2}}(kr)$ are indicated by a dash, e.g. ${}_rH'_{Tnm}$. The currents obtained by substituting (10) in (1) are represented by ${}_rI_{Snm}$, etc., those obtained from (11) by ${}_rI_{Tnm}$, etc. The various expressions are written out in table 1. All of them may be multiplied by a constant factor; the letters $A, B, \dots, G$ will be used for these constants. H_{Tnm} will be used to indicate the vector of which ${}_rH_{Tnm}$, etc., are the components, and similarly for H_{Snm} and the corresponding I 's. Sketches for $n = 1$ and 2 are given in II, figures 4*a* and *b*. The suffix m will be omitted when it is zero.

The electric fields are more complicated. From (1) and (11) the electric field corresponding to (10) is

$$E = \kappa^{-1} I_{Snm} - (\omega \wedge r) \wedge H_{Tnm} = E_{Snm} + E_{nm}^0.$$

The first term is a poloidal field similar to I_{Snm} . The second will be called E_{nm}^0 ; it is always directed radially and has a magnitude $i m \omega e^{-\sigma t} (r/k)^{\frac{1}{2}} J_{n+\frac{1}{2}}(kr) e^{im\phi} P_n^m(\theta)$. Similarly, the electric field corresponding to (11) is

$$E = \kappa^{-1} I_{Tnm} - (\omega \wedge r) \wedge H_{Snm},$$

which will be written $E = E_{Tnm} + E_{nm}^*$.

The second term is always in the meridian plane and has components

$$\begin{aligned} {}_rE_{nm}^* &= \omega e^{-\sigma t} (k^3 r)^{-\frac{1}{2}} \{kr J_{n-\frac{1}{2}}(kr) - n J_{n+\frac{1}{2}}(kr)\} e^{im\phi} \sin \theta dP_n^m(\theta)/d\theta, \\ {}_\theta E_{nm}^* &= -n(n+1) e^{-\sigma t} \omega (k^3 r)^{-\frac{1}{2}} J_{n+\frac{1}{2}}(kr) e^{im\phi} \sin \theta P_n^m(\theta). \end{aligned}$$

The electric field E_{nm}^* is the result of electromagnetic induction in the rotating sphere. It provides the essential difference between the phenomena of induction in a stationary and in a rotating sphere.

In an insulating region $\kappa = 0$ and k is zero. It is easily shown that the limits of the solutions of table 1 as $k \rightarrow 0$ are solutions of (8) with $k = 0$. The solutions of table 1 are thus still formally valid. They are not, however, convenient, as when $k \rightarrow 0$ they contain a power of k as a factor. Also the electric field $\lim_{k \rightarrow 0} (E_{Tnm} + E_{nm}^*)$ is not

purely toroidal, which is an unnecessary complication. Instead we take the solutions of table 2. These are the usual solutions of (3) and (8) with $k = 0$, and may be shown to be linear combinations of the limiting forms of table 1.† The solutions of table 2 will be distinguished by a bar over the symbol, $\bar{H}$, $\bar{E}$. The factors b^n , etc., are introduced so that the coefficient by which the $\bar{H}$ is multiplied may have the dimensions of a magnetic field. $\bar{H}'_{Snm}$ is the field due to a central magnetic multipole, $\bar{E}'_{Tnm}$ is the electric field produced if this multipole changes with time.

A zero value of k also occurs if $\sigma = i\omega m$. This arises if the field is steady and m or $\omega = 0$, or if the field rotates with the sphere. The solutions of table 2 may be used if $\sigma = \omega = 0$ even if $\kappa \neq 0$, but if ω and κ are both different from zero, table 2 may not be used even though k is zero.

† To do this it is necessary to transform E_{nm}^* by (23) below.

TABLE 1. FIELDS IN A ROTATING CONDUCTING SPHERE

Toroidal H and poloidal I, solution finite at origin

$$\begin{aligned}
{}_r H_{Tnm} &= 0 \\
{}_r H_{Tnm} &= im e^{-\sigma t} (kr)^{-\frac{1}{2}} J_{n+\frac{1}{2}}(kr) e^{im\phi} P_n^m(\theta) / \sin \theta \\
{}_r H_{Tnm} &= -e^{-\sigma t} (kr)^{-\frac{1}{2}} J_{n+\frac{1}{2}}(kr) e^{im\phi} dP_n^m(\theta) / d\theta \\
{}_r I_{Snm} &= (4\pi)^{-1} n(n+1) e^{-\sigma t} (kr^3)^{-\frac{1}{2}} J_{n+\frac{1}{2}}(kr) e^{im\phi} P_n^m(\theta) \\
{}_r I_{Snm} &= (4\pi)^{-1} e^{-\sigma t} (kr^3)^{-\frac{1}{2}} [kr J_{n-\frac{1}{2}}(kr) - n J_{n+\frac{1}{2}}(kr)] e^{im\phi} dP_n^m(\theta) / d\theta \\
{}_r I_{Snm} &= (4\pi)^{-1} m e^{-\sigma t} (kr^3)^{-\frac{1}{2}} [kr J_{n-\frac{1}{2}}(kr) - n J_{n+\frac{1}{2}}(kr)] e^{im\phi} P_n^m(\theta) / \sin \theta \\
{}_r E_{Snm} + {}_r E_{nm}^0 &= e^{-\sigma t} [n(n+1) / 4\pi\kappa + im\omega r^2] (kr^3)^{-\frac{1}{2}} J_{n+\frac{1}{2}}(kr) e^{im\phi} P_n^m(\theta) \\
{}_r E_{Snm} + {}_r E_{nm}^0 &= (4\pi\kappa)^{-1} e^{-\sigma t} (kr^3)^{-\frac{1}{2}} [kr J_{n-\frac{1}{2}}(kr) - n J_{n+\frac{1}{2}}(kr)] e^{im\phi} dP_n^m(\theta) / d\theta \\
{}_r E_{Snm} + {}_r E_{nm}^0 &= (4\pi\kappa)^{-1} im e^{-\sigma t} (kr^3)^{-\frac{1}{2}} [kr J_{n-\frac{1}{2}}(kr) - n J_{n+\frac{1}{2}}(kr)] e^{im\phi} P_n^m(\theta) / \sin \theta \\
\rho c^2 &= -(4\pi)^{-1} im\epsilon\omega e^{-\sigma t} (kr)^{-\frac{1}{2}} [kr J_{n-\frac{1}{2}}(kr) - (n-2) J_{n+\frac{1}{2}}(kr)] e^{im\phi} P_n^m(\theta)
\end{aligned}$$

*Toroidal H and poloidal I, singularity at origin*The solution is as above but with $Y_{n+\frac{1}{2}}(kr)$ in place of $J_{n+\frac{1}{2}}(kr)$.*Poloidal H and toroidal I, solution finite at origin*

$$\begin{aligned}
{}_r H_{Snm} &= n(n+1) e^{-\sigma t} (kr)^{-\frac{3}{2}} J_{n+\frac{1}{2}}(kr) e^{im\phi} P_n^m(\theta) \\
{}_r H_{Snm} &= e^{-\sigma t} (kr)^{-\frac{3}{2}} [kr J_{n-\frac{1}{2}}(kr) - n J_{n+\frac{1}{2}}(kr)] e^{im\phi} dP_n^m(\theta) / d\theta \\
{}_r H_{Snm} &= im e^{-\sigma t} (kr)^{-\frac{3}{2}} [kr J_{n-\frac{1}{2}}(kr) - n J_{n+\frac{1}{2}}(kr)] e^{im\phi} P_n^m(\theta) / \sin \theta \\
{}_r I_{Tnm} &= 0 \\
{}_r I_{Tnm} &= (4\pi)^{-1} im e^{-\sigma t} k^{\frac{1}{2}} r^{-\frac{1}{2}} J_{n+\frac{1}{2}}(kr) e^{im\phi} P_n^m(\theta) / \sin \theta \\
{}_r I_{Tnm} &= -(4\pi)^{-1} e^{-\sigma t} k^{\frac{1}{2}} r^{-\frac{1}{2}} J_{n+\frac{1}{2}}(kr) e^{im\phi} dP_n^m(\theta) / d\theta \\
{}_r E_{Tnm} + {}_r E_{nm}^* &= \omega e^{-\sigma t} (k^3 r)^{-\frac{1}{2}} [kr J_{n-\frac{1}{2}}(kr) - n J_{n+\frac{1}{2}}(kr)] e^{im\phi} \sin \theta dP_n^m(\theta) / d\theta \\
{}_r E_{Tnm} + {}_r E_{nm}^* &= e^{-\sigma t} [imk^2 / 4\pi\kappa - n(n+1)\omega \sin^2 \theta] (k^3 r)^{-\frac{1}{2}} J_{n+\frac{1}{2}}(kr) e^{im\phi} P_n^m(\theta) / \sin \theta \\
{}_r E_{Tnm} + {}_r E_{nm}^* &= -(k^{\frac{1}{2}} / 4\pi\kappa r^{\frac{1}{2}}) e^{-\sigma t} J_{n+\frac{1}{2}}(kr) e^{im\phi} dP_n^m(\theta) / d\theta \\
\rho c^2 &= -(4\pi)^{-1} \epsilon\omega e^{-\sigma t} (kr)^{-\frac{3}{2}} [\{(k^2 r^2 + 2n) J_{n+\frac{1}{2}}(kr) - 2J_{n-\frac{1}{2}}(kr)\} \sin \theta dP_n^m(\theta) / d\theta \\
&\quad + 2n(n+1) J_{n+\frac{1}{2}}(kr) \cos \theta P_n^m(\theta)] e^{im\phi}
\end{aligned}$$

*Poloidal H and toroidal I, singularity at origin*The solution is as above but with $Y_{n+\frac{1}{2}}(kr)$ in place of $J_{n+\frac{1}{2}}(kr)$.

The notation is summarized in table 3. The complexity is inherent in the problem since there are three distinct vector fields H , I and E , each of which has T and S forms. Further, there are two solutions of each form, one with and one without a singularity at the origin. Owing to the terms in E^* in the electric field these solutions are coupled, and the presence of one will usually require the presence of several others.

The solutions of table 1 involve the Bessel functions $J_{n+\frac{1}{2}}(kr)$ and $Y_{n+\frac{1}{2}}(kr)$. For real k these have been extensively tabulated by the Mathematical Tables Project of the National Bureau of Standards (1947); these tables are of little use in the investigations of the present paper, as by (9) k will be complex unless the imaginary part of $\sigma = i\omega m$. In the problems of §§ 5 to 9 σ is zero and k is proportional to $\sqrt{i}$. Tables of the functions for semi-imaginary argument were not available, and I am indebted to Mr Gellman and Mrs C. Freedman of the Computing Centre of the University of Toronto for computing tables of eight significant figures giving the real and

imaginary parts, modulus and argument of $\sqrt{(\pi/2i^{\frac{1}{2}}x)} J_{n+\frac{1}{2}}(i^{\frac{1}{2}}x)$ and $\sqrt{(\pi/2i^{\frac{1}{2}}x)} Y_{n+\frac{1}{2}}(i^{\frac{1}{2}}x)$ for $x = 0(1)10$, $n = 0, 1$ and 2 . These tables will not at present be published, but photostats are available.

TABLE 2. SOLUTIONS FOR $\kappa=0$ OR $\omega=\sigma=0$

finite at origin		singularity at origin	
<i>Toroidal H and poloidal E</i>			
$\bar{H}_{Tnm} = 0$		$\bar{H}'_{Tnm} = 0$	
$\theta \bar{H}_{Tnm} = 4\pi\kappa im(n+1)^{-1} e^{-\sigma t} (r/b)^n e^{im\phi} P_n^m(\theta)/\sin\theta$		$\theta \bar{H}'_{Tnm} = -4\pi\kappa imn^{-1} e^{-\sigma t} (b/r)^{n+1} e^{im\phi} P_n^m(\theta)/\sin\theta$	
$\phi \bar{H}_{Tnm} = -4\pi\kappa(n+1)^{-1} e^{-\sigma t} (r/b)^n e^{im\phi} dP_n^m(\theta)/d\theta$		$\phi \bar{H}'_{Tnm} = 4\pi\kappa n^{-1} e^{-\sigma t} (b/r)^{n+1} e^{im\phi} dP_n^m(\theta)/d\theta$	
$r \bar{I}_{Snm} = \kappa n e^{-\sigma t} (r^{n-1}/b^n) e^{im\phi} P_n^m(\theta)$		$r \bar{I}'_{Snm} = -\kappa(n+1) e^{-\sigma t} (b^{n+1}/r^{n+2}) e^{im\phi} P_n^m(\theta)$	
$\theta \bar{I}_{Snm} = \kappa e^{-\sigma t} (r^{n-1}/b^n) e^{im\phi} dP_n^m(\theta)/d\theta$		$\theta \bar{I}'_{Snm} = \kappa e^{-\sigma t} (b^{n+1}/r^{n+2}) e^{im\phi} dP_n^m(\theta)/d\theta$	
$\phi \bar{I}_{Snm} = i\kappa m e^{-\sigma t} (r^{n-1}/b^n) e^{im\phi} P_n^m(\theta)$		$\phi \bar{I}'_{Snm} = i\kappa m e^{-\sigma t} (b^{n+1}/r^{n+2}) e^{im\phi} P_n^m(\theta)/\sin\theta$	
$r \bar{E}_{Snm} = n e^{-\sigma t} (r^{n-1}/b^n) e^{im\phi} P_n^m(\theta)$		$r \bar{E}'_{Snm} = -(n+1) e^{-\sigma t} (b^{n+1}/r^{n+2}) e^{im\phi} P_n^m(\theta)$	
$\theta \bar{E}_{Snm} = e^{-\sigma t} (r^{n-1}/b^n) e^{im\phi} dP_n^m(\theta)/d\theta$		$\theta \bar{E}'_{Snm} = e^{-\sigma t} (b^{n+1}/r^{n+2}) e^{im\phi} dP_n^m(\theta)/d\theta$	
$\phi \bar{E}_{Snm} = im e^{-\sigma t} (r^{n-1}/b^n) e^{im\phi} P_n^m(\theta)/\sin\theta$		$\phi \bar{E}'_{Snm} = im e^{-\sigma t} (b^{n+1}/r^{n+2}) e^{im\phi} P_n^m(\theta)/\sin\theta$	
<i>Poloidal H and toroidal E</i>			
$r \bar{H}_{Snm} = n e^{-\sigma t} (r/b)^{n-1} e^{im\phi} P_n^m(\theta)$		$r \bar{H}'_{Snm} = -(n+1) e^{-\sigma t} (b/r)^{n+2} e^{im\phi} P_n^m(\theta)$	
$\theta \bar{H}_{Snm} = e^{-\sigma t} (r/b)^{n-1} e^{im\phi} dP_n^m(\theta)/d\theta$		$\theta \bar{H}'_{Snm} = e^{-\sigma t} (b/r)^{n+2} e^{im\phi} dP_n^m(\theta)/d\theta$	
$\phi \bar{H}_{Snm} = im e^{-\sigma t} (r/b)^{n-1} e^{im\phi} P_n^m(\theta)/\sin\theta$		$\phi \bar{H}'_{Snm} = im(b/r)^{n+2} e^{im\phi} P_n^m(\theta)/\sin\theta$	
$r \bar{I}_{Tnm} = \theta \bar{I}_{Tnm} = \phi \bar{I}_{Tnm} = 0$		$r \bar{I}'_{Tnm} = \theta \bar{I}'_{Tnm} = \phi \bar{I}'_{Tnm} = 0$	
$r \bar{E}_{Tnm} = 0$		$r \bar{E}'_{Tnm} = 0$	
$\theta \bar{E}_{Tnm} = im\sigma(n+1)^{-1} (r^n/b^{n-1}) e^{im\phi} P_n^m(\theta)/\sin\theta$		$\theta \bar{E}'_{Tnm} = -im\sigma n^{-1} (b^{n+2}/r^{n+1}) e^{im\phi} P_n^m(\theta)/\sin\theta$	
$\phi \bar{E}_{Tnm} = -\sigma(n+1)^{-1} (r^n/b^{n-1}) e^{im\phi} dP_n^m(\theta)/d\theta$		$\phi \bar{E}'_{Tnm} = \sigma n^{-1} (b^{n+2}/r^{n+1}) e^{im\phi} dP_n^m(\theta)/d\theta$	

TABLE 3. SUMMARY OF NOTATION

	no singularity at origin				singularity at origin			
	toroidal		poloidal		toroidal		poloidal	
	$k \neq 0$	$k = 0$	$k \neq 0$	$k = 0$	$k \neq 0$	$k = 0$	$k \neq 0$	$k = 0$
magnetic field:								
vector	H_{Tnm}	$\bar{H}_{Tnm}$	H_{Snm}	$\bar{H}_{Snm}$	H'_{Tnm}	$\bar{H}'_{Tnm}$	H'_{Snm}	$\bar{H}'_{Snm}$
r component	$r H_{Tnm}$	$r \bar{H}_{Tnm}$	$r H_{Snm}$	$r \bar{H}_{Snm}$	$r H'_{Tnm}$	$r \bar{H}'_{Tnm}$	$r H'_{Snm}$	$r \bar{H}'_{Snm}$
θ component	θH_{Tnm}	$\theta \bar{H}_{Tnm}$	θH_{Snm}	$\theta \bar{H}_{Snm}$	$\theta H'_{Tnm}$	$\theta \bar{H}'_{Tnm}$	$\theta H'_{Snm}$	$\theta \bar{H}'_{Snm}$
ϕ component	ϕH_{Tnm}	$\phi \bar{H}_{Tnm}$	ϕH_{Snm}	$\phi \bar{H}_{Snm}$	$\phi H'_{Tnm}$	$\phi \bar{H}'_{Tnm}$	$\phi H'_{Snm}$	$\phi \bar{H}'_{Snm}$
electric field:								
vector	E_{Tnm}	$\bar{E}_{Tnm}$	E_{Snm}	$\bar{E}_{Snm}$	E'_{Tnm}	$\bar{E}'_{Tnm}$	E'_{Snm}	$\bar{E}'_{Snm}$
current:								
vector	I_{Tnm}	$\bar{I}_{Tnm}$	I_{Snm}	$\bar{I}_{Snm}$	I'_{Tnm}	$\bar{I}'_{Tnm}$	I'_{Snm}	$\bar{I}'_{Snm}$

4. THE BOUNDARY CONDITIONS

For any values of n and m tables 1 and 2 give four complex solutions of the equations (1) to (4). These solutions must be combined to satisfy the boundary conditions. At the outer surface of the shell I_r is zero and H_r , H_θ , H_ϕ , E_θ and E_ϕ are continuous; at the interface between the rotating sphere and the shell all these quantities are continuous (Tolman 1934, pp. 110–112). The fields and current will either vanish at a great distance from the sphere or their behaviour will be specified.

We consider separately induction by a steady external field (§§ 5 and 6), induction by a constant multipole at the origin (§ 7), induction by the field due to a source of current (§ 8), and the freely decaying modes (§ 11). Induction by a variable field will not be considered, though it presents no difficulty of principle.

5. ROTATING SPHERE IN A UNIFORM FIELD

The process of solution is best illustrated by a special case. Consider a conducting sphere of radius b rotating in an insulating medium, and suppose a uniform magnetic field to be applied at right angles to the axis.

A uniform field along $\phi = 0$ is the real part of $F\bar{H}_{S11}$, where F is the magnitude of the field. The boundary conditions for H and I can be satisfied if an additional field $F'\bar{H}'_{S11}$ (a dipole field) is introduced outside the sphere and a field $A\bar{H}_{S11}$ inside it. These fields give:

inside	outside
$H_r = 2A(kr)^{-\frac{3}{2}}J_{\frac{3}{2}}(kr)e^{i\phi}\sin\theta$	$H_r = (F - 2F'b^3/r^3)e^{i\phi}\sin\theta$
$H_\theta = -A\{J_{\frac{3}{2}}(kr) - krJ_{\frac{5}{2}}(kr)\}(kr)^{-\frac{3}{2}}e^{i\phi}\cos\theta$	$H_\theta = (F + F'b^3/r^3)e^{i\phi}\cos\theta$
$H_\phi = -iA\{J_{\frac{3}{2}}(kr) - krJ_{\frac{5}{2}}(kr)\}(kr)^{-\frac{3}{2}}e^{i\phi}$	$H_\phi = i(F + F'b^3/r^3)e^{i\phi}$
$I_r = 0$	$I_r = 0$
$I_\theta = iA(k^{\frac{3}{2}}/4\pi r^{\frac{3}{2}})J_{\frac{3}{2}}(kr)e^{i\phi}$	$I_\theta = 0$
$I_\phi = -A(k^{\frac{3}{2}}/4\pi r^{\frac{3}{2}})J_{\frac{3}{2}}(kr)e^{i\phi}\cos\theta$	$I_\phi = 0$
$E_r = -A\omega k^{-\frac{3}{2}}r^{-\frac{3}{2}}\{J_{\frac{3}{2}}(kr) - krJ_{\frac{5}{2}}(kr)\}e^{i\phi}\sin\theta\cos\theta$	$E_r = 0$
$E_\theta = A\omega k^{-\frac{3}{2}}r^{-\frac{3}{2}}J_{\frac{3}{2}}(kr)e^{i\phi}\cos 2\theta$	$E_\theta = 0$
$E_\phi = iA\omega k^{-\frac{3}{2}}r^{-\frac{3}{2}}J_{\frac{3}{2}}(kr)e^{i\phi}\cos\theta$	$E_\phi = 0$

where $k = \sqrt{(-4\pi\kappa i\omega)}$.

The boundary conditions for H are satisfied if

$$\left. \begin{aligned} A &= \frac{3}{2}k^{\frac{3}{2}}b^{\frac{3}{2}}F/J_{\frac{3}{2}}(kb), \\ F' &= \frac{1}{2}F\{1 - 3J_{\frac{3}{2}}(kb)/kbJ_{\frac{5}{2}}(kb)\} = -\frac{1}{2}FJ_{\frac{3}{2}}(kb)/J_{\frac{5}{2}}(kb). \end{aligned} \right\} \quad (13)$$

These values of A and F' will not make E_θ and E_ϕ continuous at $r = b$. What is needed is an electric field outside the sphere which vanishes at infinity, and whose θ and ϕ components vary as $e^{i\phi}\cos 2\theta$ and $e^{i\phi}\cos\theta$. From table 2 it is found that this field is $\bar{E}'_{S21}$, and that it requires no addition to the magnetic field (since $\kappa = 0$ outside the sphere). This step is the crucial one in the solution of most of the problems of the present paper. The part E^* of the electric field in the moving medium can only be balanced by introducing fields with an n different from that occurring in H . The problem of this section is specially simple, as only one such electric field is needed, and the corresponding magnetic field is zero.

If the additional field is $G'\bar{E}'_{S21}$, its components are†

$$\begin{aligned} E_r &= -9G'b^3r^{-4}e^{i\phi}\sin\theta\cos\theta, \\ E_\theta &= 3G'b^3r^{-4}e^{i\phi}\cos 2\theta, \\ E_\phi &= 3G'ib^3r^{-4}e^{i\phi}\cos\theta. \end{aligned}$$

† The associated Legendre function $P_n^m(\theta)$ is taken as $\sin^m\theta\partial^m P_n(\theta)/\partial(\cos\theta)^m$; this differs from the partially normalized functions usual in numerical work in terrestrial magnetism, but is more convenient in an analytical problem. $P_n^m(\theta)$ has been written for typographical simplicity where $P_n^m(\cos\theta)$ is more usual.

The boundary conditions are satisfied if

$$G' = \frac{\omega b J_{\frac{1}{2}}(kb)}{2k J_{\frac{3}{2}}(kb)} F.$$

The external magnetic field consists of the constant field F , and of the real part of the term in F' . If

$$F' = -Mb^{-3}e^{i\zeta}, \quad (14)$$

where M and ζ are real, then the real part of the field is

$$H_r = 2Mr^{-3} \cos(\phi - \zeta) \sin \theta, \quad H_\theta = -Mr^{-3} \cos(\phi - \zeta) \cos \theta.$$

The field is therefore similar to that of a dipole of moment M with its axis at right angles to the axis of rotation and directed along $\phi = \zeta$.

$$\text{From (13) and (14)} \quad Me^{-i\zeta} = \frac{1}{2} Fb^3 J_{\frac{1}{2}}(kb) / J_{\frac{3}{2}}(kb). \quad (15)$$

This may be expressed in terms of elementary functions and gives

$$M = Fb^3 |(\alpha^2 + \beta^2)^{\frac{1}{2}}|, \quad \zeta = \pi - \tan \beta / \alpha,$$

$$\text{where} \quad \alpha = \frac{1}{2} - \frac{3}{2x} \frac{\sinh x - \sin x}{\cosh x - \cos x}, \quad \beta = \frac{3}{x^2} - \frac{3}{2x} \frac{\sinh x + \sin x}{\cosh x - \cos x},$$

$$\text{and} \quad x = \sqrt{(8\pi\kappa\omega b^2)}.$$

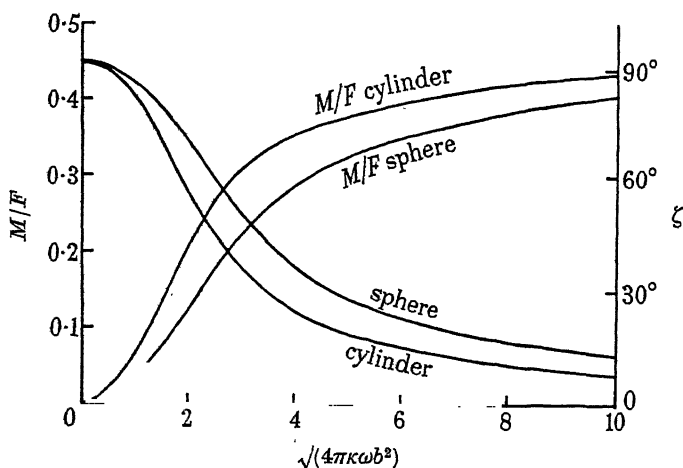


FIGURE 1. Dipole moment of a sphere or a cylinder rotating in a field at right angles to the axis of rotation. The ordinate is M/Fb^3 for the sphere and M/Fb^3 for the cylinder.

M and ζ have been computed from these expressions and are given in figure 1. For small x the moment is at right angles to the initial field and of magnitude $2\pi\kappa\omega b^5 F/15$. As the velocity of rotation increases the moment swings round and ultimately becomes anti-parallel to the initial field. Its magnitude tends to the limit $\frac{1}{2}Fb^3$, which is

equal to that of a perfectly diamagnetic sphere.† For small and large x the following expressions may be used:

$$\left. \begin{aligned} M/Fb^3 &= x^2/60 - 11x^6/529200 \\ \tan \zeta &= 21/x^2 + 247x^2/660 \end{aligned} \right\} \text{ for small } x,$$

$$\left. \begin{aligned} M/Fb^3 &= \frac{1}{2} - \frac{3}{2x} + \frac{9}{4x^2} - \frac{9}{4x^3} \\ \tan \zeta &= \frac{3(x-2)}{x(x-3)} \end{aligned} \right\} \text{ for large } x.$$

The electric field outside the sphere is similar to that due to an electric quadrupole located at the centre of the sphere. It is similar to the S_2^c field sketched in figure 4*b* of II; for slow rotation the plane of the quadrupole is $\phi = 0$, for rapid rotation it is $\phi = \frac{1}{4}\pi$.

For slow rotation the magnetic field inside the sphere is simply the incident constant field. For fast rotation it tends to zero and the sphere behaves like a perfect diamagnetic. The radial variation of the maximum value of H_r is shown in figure 2 for various values of $4\pi\kappa\omega b^2$; for fast rotation the field behaves like

$$r^{-\frac{1}{2}} \exp [-(2\pi\kappa\omega)^{\frac{1}{2}}(b-r)],$$

and is small except within a distance $(2\pi\kappa\omega)^{-\frac{1}{2}}$ from the surface.

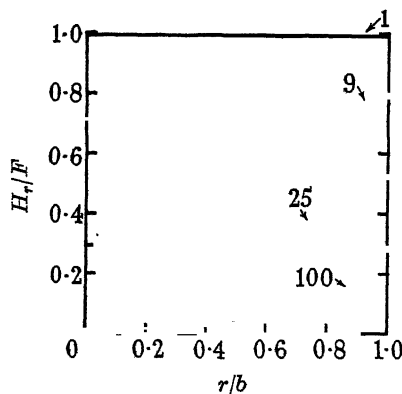


FIGURE 2. Radial variation of the field within a sphere rotating in a field at right angles to the axis of rotation. The numbers on the curves give the values of $4\pi\kappa\omega b^2$.

There is a discontinuity in the radial electric field at the surface of the sphere. This implies a charge of order $\omega b F/c^2$ per unit area. This is balanced by a volume distribution of charge given by (4). The motion of these charges would produce a magnetic field smaller than F by a factor of the order $(\omega b/c)$. This is negligible if the velocities involved are much less than that of light. The omission of such terms in (1) is therefore justified. The terms containing the displacement current are also of the same order of magnitude.

† These limiting forms agree with those obtained by Thomson (1893). Different results, which are believed to be incorrect, have been obtained by Gans (1921) and quoted by Valentiner (1927, p. 360).

Similar calculations have been made for a cylinder of radius b rotating with an angular velocity ω in a field F at right angles to its axis. The moment M per unit length and the angle that the direction of the moment makes with the field F are found to be given by

$$M e^{-i\zeta} = b^2 J_2(y) / J_0(y), \quad \text{where} \quad y^2 = 4\pi i \kappa \omega b^2. \quad (16)$$

This is the analogue of (15). For slow rotation,

$$M = \frac{1}{4} \pi \kappa \omega b^4, \quad \zeta = \frac{1}{2} \pi.$$

For fast rotation M tends to the constant value $\frac{1}{2} b^2 F$. As with the sphere it becomes antiparallel to F and produces a zero resultant field within the cylinder. M/b^2 and ζ have been computed from (16) as functions of $4\pi \kappa \omega b^2$; the results, which have been used in I, are given in figure 1.

6. INDUCTION BY A FIXED EXTERNAL FIELD

We now generalize the results of the last section by supposing the rotating sphere to be surrounded by a stationary shell of radius a . The field at a great distance outside the shell will be supposed to be $F \bar{H}_{Snm}$ and to be constant in time. The field of the last section is $F \bar{H}_{S11}$, and is therefore a special case of this. The magnetic field in the rotating sphere must contain $A H_{Snm}$, that in the shell $B \bar{H}_{Snm}$ and $B' \bar{H}'_{Snm}$, and that outside the shell $F' \bar{H}'_{Snm}$, where A , B , B' and F' are constants. The boundary conditions for the magnetic field at $r = a$ give

$$B \bar{H}_{Snm} + B' \bar{H}'_{Snm} = F \bar{H}_{Snm} + F' \bar{H}'_{Snm}, \quad (17)$$

which can only be satisfied if $B = F$ and $B' = F'$. The conditions at $r = b$ then give

$$A H_{Snm} = F \bar{H}_{Snm} + F' \bar{H}'_{Snm}. \quad (18)$$

Using tables 1 and 2 these are found to be satisfied if

$$A = \frac{2n+1}{n+1} \frac{(kb)^{\frac{1}{2}}}{J_{n-\frac{1}{2}}(kb)} F, \quad (19)$$

$$F' = -\frac{n}{n+1} F J_{n+\frac{1}{2}}(kb) / J_{n-\frac{1}{2}}(kb). \quad (20)$$

These relations do not contain a . It will be shown below that they are not disturbed by the additional fields necessary to satisfy the remaining boundary conditions. We have therefore the important result that the external magnetic field produced by the rotation of a sphere in an arbitrary external magnetic field is not influenced by the presence of a concentric conducting shell (compare (13) and (20)).

The fields considered above do not satisfy the boundary conditions for the electric field at $r = b$. In the shell the electric field corresponding to $\bar{H}_{Snm}$ and $\bar{H}'_{Snm}$ is zero, whilst in the rotating sphere at $r = b$,

$$E_\theta = A \omega \{m^2 - n(n+1) \sin^2 \theta\} (k^3 b)^{-\frac{1}{2}} J_{n+\frac{1}{2}}(kb) e^{im\phi} P_n^m(\theta) / \sin \theta, \quad (21)$$

$$E_\phi = i A \omega m (k^3 b)^{-\frac{1}{2}} J_{n+\frac{1}{2}}(kb) e^{im\phi} dP_n^m(\theta) / d\theta. \quad (22)$$

There are no solutions in table 2 that will match these as they stand, but it may be shown that

$$\sin \theta dP_n^m(\theta) / d\theta = [n(n-m+1) P_{n+1}^m(\theta) - (n+1)(n+m) P_{n-1}^m(\theta)] / (2n+1), \quad (23)$$

where $P_{n-1}^m(\theta)$ is to be put equal to zero if $n = m$. If this is substituted in (22) it gives

$$E_\phi = A \frac{i\omega m}{2n+1} \frac{J_{n+\frac{1}{2}}(kb)}{k^{\frac{1}{2}}b^{\frac{1}{2}}} e^{im\phi} [n(n-m+1)P_{n+1}^m(\theta) - (n+1)(n+m)P_{n-1}^m(\theta)]/\sin\theta.$$

This can be matched by fields $\bar{E}_{S(n\pm 1)m}$ and $\bar{E}'_{S(n\pm 1)m}$ in the shell. If these fields are introduced the corresponding magnetic fields from table 2 must also be included.

Inside the rotating sphere $E_{S(n\pm 1)m}$ and the corresponding magnetic fields are also to be expected, but not the $E'_{S(n\pm 1)m}$, as they have a singularity at the origin. Outside the shell $\bar{E}'_{S(n\pm 1)m}$ will occur; these imply no additional magnetic field. The complete set of fields is

$$\left. \begin{aligned} H &= AH_{Snm} + D_{n+1}H_{T(n+1)m} + D_{n-1}H_{T(n-1)m}, \\ I &= AI_{Tnm} + D_{n+1}I_{S(n+1)m} + D_{n-1}I_{S(n-1)m}, \\ E &= A(E_{Tnm} + E_{nm}^*) + D_{n+1}(E_{S(n+1)m} + E_{(n+1)m}^o) \\ &\quad + D_{n-1}(E_{S(n-1)m} + E_{(n-1)m}^o), \end{aligned} \right\} \begin{array}{l} \text{in the} \\ \text{sphere,} \end{array}$$

$$\left. \begin{aligned} H &= B\bar{H}_{Snm} + B'\bar{H}'_{Snm} + C_{n+1}\bar{H}_{T(n+1)m} + C'_{n+1}\bar{H}'_{T(n+1)m} + C_{n-1}\bar{H}_{T(n-1)m} \\ &\quad + C'_{n-1}\bar{H}'_{T(n-1)m}, \\ I &= C_{n+1}\bar{I}_{S(n+1)m} + C'_{n+1}\bar{I}'_{S(n+1)m} + C_{n-1}\bar{I}_{S(n-1)m} + C'_{n-1}\bar{I}'_{S(n-1)m}, \\ E &= C_{n+1}\bar{E}_{S(n+1)m} + C'_{n+1}\bar{E}'_{S(n+1)m} + C_{n-1}\bar{E}_{S(n-1)m} + C'_{n-1}\bar{E}'_{S(n-1)m}, \end{aligned} \right\} \begin{array}{l} \text{in the} \\ \text{shell,} \end{array} \quad (24)$$

$$\left. \begin{aligned} H &= F\bar{H}_{Snm} + F'\bar{H}'_{Snm}, \\ I &= 0, \\ E &= G'_{n+1}\bar{E}'_{S(n+1)m} + G'_{n-1}\bar{E}'_{S(n-1)m}, \end{aligned} \right\} \text{outside,}$$

where the C 's, D 's and G 's are constants. The terms in $(n-1)$ will not be required when $n = m$. When $n \neq 1$, $m = 0$ they must also be omitted as they give a net charge on the sphere. They may then be regarded as cancelled by equal and opposite terms from table 1 with $n = 0$. As the new terms all have $n \pm 1$ where (17) and (18) have n , the conditions satisfied by A , B , B' and F' are not affected and (19) and (20) are still true. The boundary conditions for I_r or H_θ give $C_{n\pm 1}$ and $D_{n\pm 1}$ in terms of $C'_{n\pm 1}$. Those for E_ϕ at $r = a$ give $G'_{n\pm 1}$ in terms of $C'_{n\pm 1}$. The E_ϕ condition at $r = b$ when transformed by (23) then gives $C'_{n\pm 1}$, and thus all the other constants, in terms of A which is related to the applied field F by (19). The E_θ condition at $r = b$ may be shown to be the differential with respect to θ of the E_ϕ condition. It is therefore necessarily satisfied if the latter is.

The results are

$$C'_{n+1} = \frac{n+1}{n+2} \left(\frac{a}{b}\right)^{2n+3} C_{n+1} = \frac{n+1}{2n+3} G'_{n+1} = \frac{n(n-m+1)\omega F J_{n+\frac{1}{2}}(kb)}{k^2 J_{n-\frac{1}{2}}(kb) \{1 + (b/a)^{2n+3} J_{n+\frac{1}{2}}(kb)/J_{n-\frac{1}{2}}(kb)\}},$$

$$C'_{n-1} = \frac{n-1}{n} \left(\frac{a}{b}\right)^{2n-1} C_{n-1} = \frac{n-1}{2n-1} G'_{n-1} = -\frac{(n-1)(n+m)\omega F J_{n+\frac{1}{2}}(kb)}{k^2 J_{n-\frac{1}{2}}(kb) \{1 + (b/a)^{2n-1} J_{n+\frac{1}{2}}(kb)/J_{n-\frac{1}{2}}(kb)\}},$$

$$D_{n+1} = -\frac{4\pi\kappa n(n-m+1)\omega F b^{\frac{1}{2}} \{1 - (b/a)^{2n+3}\}}{(n+1)k^{\frac{1}{2}} J_{n-\frac{1}{2}}(kb) \{1 + (b/a)^{2n+3} J_{n+\frac{1}{2}}(kb)/J_{n-\frac{1}{2}}(kb)\}},$$

$$D_{n-1} = \frac{4\pi\kappa(n+m)\omega F b^{\frac{1}{2}} J_{n+\frac{1}{2}}(kb) \{1 - (b/a)^{2n+1}\}}{k^{\frac{1}{2}} J_{n-\frac{1}{2}}(kb) J_{n-\frac{1}{2}}(kb) \{1 + (b/a)^{2n-1} J_{n+\frac{1}{2}}(kb)/J_{n-\frac{1}{2}}(kb)\}}.$$

These coefficients can be substituted in (24) to give the electric and magnetic fields and currents. The expressions are rather complicated; the feature of most interest is the toroidal magnetic field $D_{n\pm 1}H_{T(n\pm 1)m}$ in the sphere and

$$C_{n\pm 1}H_{T(n\pm 1)m} + C'_{n\pm 1}H_{T(n\pm 1)m}$$

in the shell. This field lies entirely within the conductor and has no radial component. Sketches of typical forms have been given in figure 4*a* of II. The behaviour of the fields for $m = 0$ is different from that when m is not equal to 0.

(i) $m = 0$. When $m = 0$, k is zero, and all the Bessel functions may be replaced by the first term in their ascending series for all speeds of rotation. This produces a considerable simplification. The toroidal field is entirely in the ϕ direction and is

$$H_{\phi} = \frac{4\pi\kappa F\omega n b^{n+4}}{(2n+3)(2n+1)r^{n+2}} \left[\left(1 - \frac{r^{2n+3}}{a^{2n+3}}\right) \frac{dP_{n+1}}{d\theta} - \frac{2n+3}{2n-1} \frac{r^2}{b^2} \left(1 - \frac{r^{2n-1}}{a^{2n-1}}\right) \frac{dP_{n-1}}{d\theta} \right]$$

in the shell. In the sphere it is

$$H_{\phi} = \frac{4\pi\kappa F\omega n r^{n-1}}{(2n+3)(2n+1)b^{n-3}} \left[\frac{r^2}{b^2} \left(1 - \frac{b^{2n+3}}{a^{2n+3}}\right) \frac{dP_{n+1}}{d\theta} - \frac{2n+3}{2n-1} \left(1 - \frac{b^{2n-1}}{a^{2n-1}}\right) \frac{dP_{n-1}}{d\theta} \right].$$

The field is proportional to the rate of rotation and increases indefinitely as ω is increased. Owing to the high powers to which b/a and r/a are raised, quite a thin shell will produce a close approximation to the field that occurs when a is indefinitely great.

(ii) $n = 1$, $m = 0$. The applied field, F , is uniform and parallel to the axis of rotation. k is then zero, and so is D_2 , which contains k as a factor. As this k cancels when it is substituted in the expressions for the field the toroidal field is finite in spite of the vanishing of D_2 .

F' is found from (20) to vanish. There is therefore no external magnetic field except the initial field F . There is an external electric field in meridian planes giving

$$E_r = -\frac{1}{2}\omega F(3\cos^2\theta - 1)b^5/r^4, \quad E_{\theta} = -\frac{1}{2}\omega F \sin 2\theta b^5/r^4,$$

which is the field of an electric quadrupole with all its poles on the axis.

The magnetic field in the shell consists of the initial field F and the field produced by the induced currents. The latter is entirely in the ϕ direction. Its lines of force are circles of latitude. It is given by

$$H_{\phi} = -\frac{2}{3}\pi\kappa\omega F b^5 r^2 (1/r^5 - 1/a^5) \sin 2\theta. \quad (25a)$$

This has its maximum value at $r = b$ and $\theta = \frac{1}{4}\pi$,

$$H_{\phi}(\text{max.}) = -\frac{2}{3}\pi\kappa\omega F b^2 (1 - b^5/a^5). \quad (25b)$$

This is the result on which § 2 of II is based. The longest time of relaxation τ_0 of a freely decaying current system in a stationary sphere of radius b is $4b^2\kappa/\pi$ (see § 11). The maximum field for $a \gg b$ is therefore

$$H_{\phi}(\text{max.}) = \frac{\pi^3 \tau_0}{5 \tau_1} F,$$

where τ_1 is the time for one revolution. The factor by which the toroidal field exceeds the inducing field is thus about 6.2 times the number of rotations made in the time in which a freely decaying field would fall to $1/e$ of its initial value.

The current in the shell producing the toroidal field flows in meridian planes and has components,

$$I_r = \frac{1}{5}\kappa\omega F r b^5 (1/r^5 - 1/a^5) (1 - 3 \cos^2 \theta), \quad I_\theta = -\frac{1}{5}\kappa\omega F r b^5 (1/r^5 + 3/2a^5) \sin 2\theta.$$

The corresponding electric field is $1/\kappa$ times this. Inside the rotating sphere the magnetic field produced by the induced currents resembles that in the shell and is given by

$$H_\phi = -\frac{2}{5}\pi\kappa\omega F r^2 (1 - b^5/a^5) \sin 2\theta.$$

The induced current and the electric field in the sphere are in meridian planes and are given by

$$\left. \begin{aligned} I_r &= \frac{1}{5}\kappa\omega F r (1 - b^5/a^5) (1 - 3 \cos^2 \theta), \\ I_\theta &= \frac{3}{10}\kappa\omega F r (1 - b^5/a^5) \sin 2\theta, \\ E_r &= \omega F r [\frac{1}{5}(1 - 3 \cos^2 \theta) (1 - b^5/a^5) - \sin^2 \theta], \\ E_\theta &= -\frac{1}{5}\omega F r (1 + 3b^5/2a^5) \sin 2\theta. \end{aligned} \right\} \quad (26)$$

The volume density of charge in the rotating sphere is found from (4) to be constant and equal to $-\epsilon\omega F/2\pi c^2$. The total charge is $-2\epsilon\omega F b^3/3c^2$. On the surface of the rotating sphere there is a superficial charge density of

$$(\epsilon\omega F b/4\pi c^2) \sin^2 \theta.$$

When integrated over the surface of the sphere this balances the charge distributed through the volume.

There is no volume charge in the shell and a superficial charge density of

$$-(\epsilon\omega F b/8\pi c^2) (3 \cos^2 \theta - 1)$$

on its outer surface at $r = a$. This gives no net charge when integrated over the sphere.

The total energy in the toroidal magnetic field may be found by integrating $H^2/8\pi$ through the sphere and the shell. When $a \gg b$ the result is

$$\begin{aligned} \text{energy in sphere} &= 4b^3 H_\phi^2 (\text{max.})/105, \\ \text{energy in shell} &= 4b^3 H_\phi^2 (\text{max.})/45, \\ \text{total energy} &= 8b^3 H_\phi^2 (\text{max.})/63. \end{aligned}$$

Thus only 30 % of the energy is in the sphere and the rest in the shell.

The lines of electric force and current flow are easily found from the expressions given above. They have been used to compute the example shown in figure 3, where $a = 1.5b$. In figure 3a the four lines of current flow may be considered as sections of surfaces obtained by revolving the figure about its axis. These surfaces divide the hemisphere into five parts, each of which carries an equal total current. The lines of electric force in figure 3b have been chosen so that each of the five volumes carries an equal electric flux across the surface of a sphere of radius $< b$. Owing to the distribution of charge through the volume, the flux is not the same for all spheres. The singularity at the centre is only apparent as the electric field is zero there.

Figure 3c shows contours of equal H_ϕ . Figure 4 shows the radial variation of H_ϕ . In these figures F has been taken as in a negative direction so that the part of the sphere shown corresponds to the northern hemisphere of the earth where the vertical component of the field is directed downwards.

(iii) $m \neq 0$. The main interest is in the behaviour of the toroidal magnetic field as ω increases. For sufficiently large ω the first term of the terminating series for the Bessel functions may be used. As before the exact value of a is of little importance

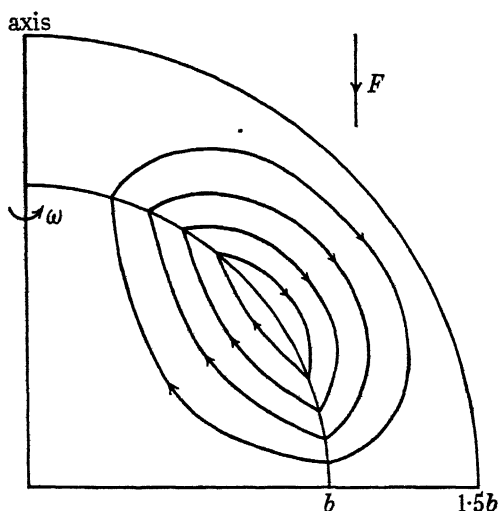


FIGURE 3a. Lines of current flow producing the toroidal field.

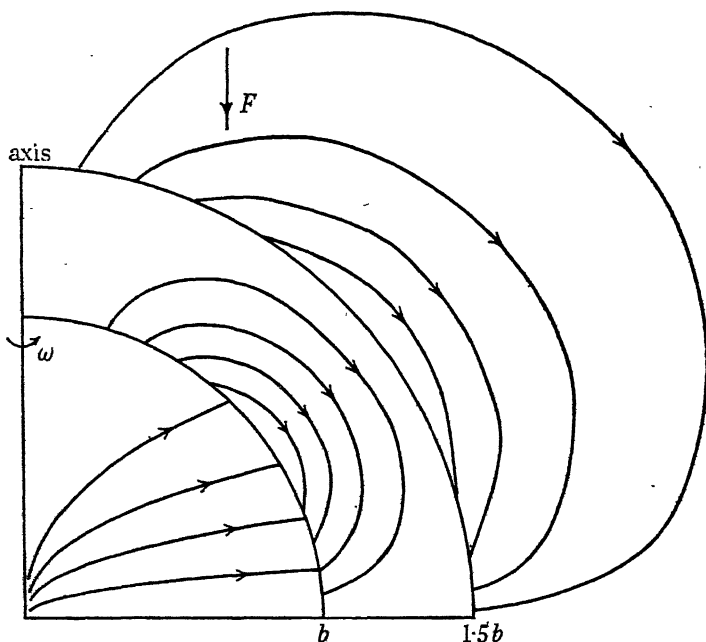


FIGURE 3b. Lines of electric force associated with the toroidal field.

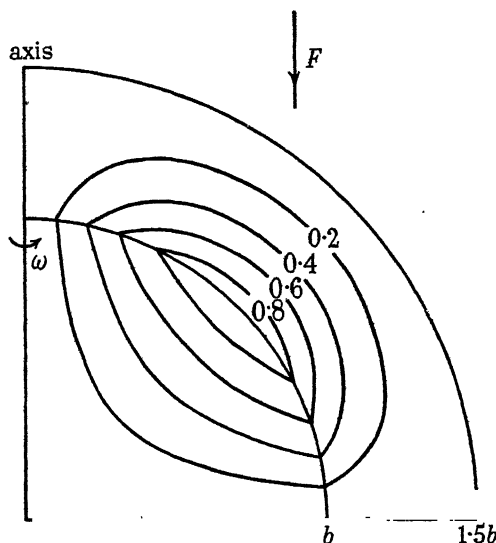


FIGURE 3c. Contours of the toroidal field. The figures on the curves give H_ϕ as a fraction of its maximum value.

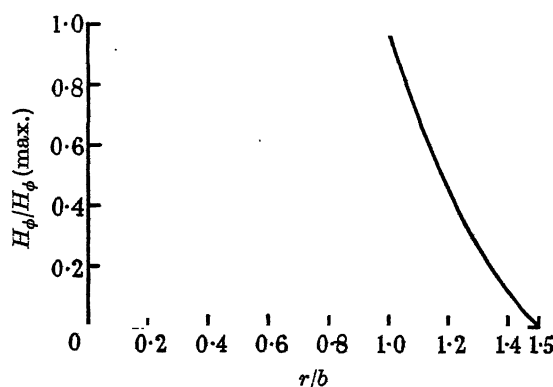


FIGURE 4. Radial variation of the toroidal field.

so long as it is substantially larger than b . At $r=b$, where the field is near its maximum, and for $a \gg b$ the toroidal field becomes after some transformation,

$$H_\theta = -\frac{2n+1}{n+1} F \cos m\phi \frac{dP_n^m(\theta)}{d\theta},$$

$$H_\phi = \frac{2n+1}{m(n+1)} F \{n(n+1) \sin^2 \theta - m^2\} P_n^m(\theta) \sin m\phi / \sin \theta.$$

This does not contain ω or κ , and shows that the field tends to a finite limit as the speed of rotation is increased. As the inducing field has components

$$H_r = nF \cos m\phi P_n^m(\theta), \quad H_\theta = F \cos m\phi dP_n^m(\theta)/d\theta, \quad H_\phi = mF \sin m\phi P_n^m(\theta)/\sin \theta,$$

the induced field exceeds the corresponding component of the inducing field by at most a factor $(2n+1)n$.

The magnetic field within the sphere due to the induced multipole is AH_{Snm} . This may be shown from (19) to vanish for fast rotation in a way qualitatively similar to the field of § 5 figure 2.

The induced toroidal field with $m \neq 0$ differs markedly from that when $m = 0$, for the latter increases without limit as the speed of rotation is increased. The physical reason for this is that when $m \neq 0$ the induced multipole opposes the initial field and weakens it more and more as the speed of rotation increases. For $m = 0$ there is no induced multipole.

7. INDUCTION BY A CENTRAL MULTIPOLE

The field due to a magnetic multipole at the centre of the rotating sphere is $A'H'_{Snm}$, where A' is a constant. If the multipole is stationary σ is zero; k^2 is then equal to $-4\pi i k \omega m$ in the sphere, and is zero in the shell. If the multipole rotates with the sphere $\sigma = i \omega m$ and k^2 is $4\pi i k \omega m$ in the shell and zero in the sphere.

The problem of the stationary multipole can be solved in the same way as that of § 6. It is only necessary to add $A'H'_{Snm}$ and the corresponding electric field and current to the fields in the sphere given by (24), and to omit the term $F\bar{H}_{Snm}$. As before the solution contains poloidal and toroidal parts. With the notation of (24) the coefficients of the poloidal part are,

$$A = -A'Y_{n-\frac{1}{2}}(kb)/J_{n-\frac{1}{2}}(kb), \quad F' = B' = (2/\pi)nA'/(kb)^{\frac{1}{2}}J_{n-\frac{1}{2}}(kb).$$

B is zero and there is therefore no $\bar{H}_{Snm}$ field in the shell.

If $n = 1$ and the sphere is rotating slowly $A'H'_{S11}$ is the field due to a dipole of moment $M = -\sqrt{2}A'/\sqrt{\pi}k^3$ directed along $\theta = \frac{1}{2}\pi$, $\phi = 0$. For any rate of rotation the external field, $F'\bar{H}'_{S11}$, is that due to a dipole of moment $-F'b^3$ given by

$$-F'b^3 = Mkb/\sin kb.$$

For slow rotation this is equal to M as it should be. For $m = 0$, that is, when the axis of the dipole lies along the axis of rotation, k is zero for all speeds of rotation and the external field is unaffected by the rotation. If $m = 1$, k is semi-imaginary and the external field tends to zero for fast rotation, the moment being $2M\rho^{\frac{1}{2}}\exp(-\rho^{\frac{1}{2}}/\sqrt{2})$, where $\rho = 4\pi k \omega b^2$. The field is thus very effectively prevented from emerging if the rotation is fast enough to make ρ large. For large ρ the direction of the external dipole is at 45° to that of M .

The toroidal fields in the shell and in the sphere are very similar to those found in § 6. The fields may be found by substituting

$$-2A'(n+1)/\pi(2n+1)(kb)^{\frac{1}{2}}J_{n+\frac{1}{2}}(kb) \quad (27)$$

for F in the previous expressions for the C 's, D 's, E 's and G 's. This factor leaves the form of the fields and currents unchanged and only affects their variation with the angular velocity.

If $m = 0$, k is zero and the factor (27) reduces to H_0/H_1 , where H_0 and H_1 are the fields at $r = b$ produced by the centred dipole and by the external field of § 6. Thus if $m = 0$ the toroidal field induced by a central magnetic singularity is identical with that

induced by an external field that gives the same field at the surface. This result shows the toroidal field for $m = 0$ to be insensitive to the radial variation of the inducing field, and to depend only on its value near $r = b$. This result has been used in II.

8. INDUCTION WITH CURRENT SOURCES

The expressions of §§ 6 and 7 enable the currents and fields produced by induction by an assumed poloidal magnetic field to be found; for any such field can be expressed as a sum of terms like H_{Snm} and H'_{Snm} within the sphere, and $\bar{H}_{Snm}$ and $\bar{H}'_{Snm}$ outside it.

In this section we consider problems in which the inducing magnetic field is toroidal. Taken literally these solutions are of little practical importance, since they imply current sources either at the origin or at infinity. Similar solutions are, however, possible in which the toroidal field is produced by motions more complicated than the rotation considered in this paper. The interaction with a rotation of fields produced in this way is of importance in § 8 of II, and it is believed that the present investigation throws some light on this.

In view of the insensitivity of the solutions of §§ 6 and 7 to the outer radius of the shell, a will be taken as infinite in this section. If $a = b$ there is no conducting shell and no external magnetic field.

Current source at centre

A current source at the centre of the rotating sphere may be represented by a term $D'H'_{Tnm}$ in the magnetic field and corresponding terms in E and I . The boundary conditions can be satisfied by the addition of a field DH_{Tnm} in the sphere and a field $C'H'_{Tnm} + CH_{Tnm}$ outside it. If the source rotates with the sphere, $\sigma = i\omega m$ and $k = 0$ in the sphere. Outside the sphere $k^2 = 4\pi i k \omega m$. If $\omega = 0$ the θ and ϕ components of the field at $r = b$ are proportional to the real parts of

$$im e^{i(m\phi - \omega t)} P_n^m(\theta)/n \sin \theta, \quad \text{and} \quad -n^{-1} e^{i(m\phi - \omega t)} dP_n^m(\theta)/d\theta.$$

Rotation changes this by a factor

$$(2n+1) H_{n+\frac{1}{2}}^{(1)}(kb)/kb H_{n+\frac{1}{2}}^{(1)}(kb),$$

where $H_{n+\frac{1}{2}}^{(1)}$ is the Hankel function $J_{n+\frac{1}{2}} + iY_{n+\frac{1}{2}}$. For slow rotation this is unity. For rapid rotation it is $(2n+1) e^{i\pi/2} / \sqrt{(4\pi k \omega b^2)}$. The field therefore rotates with the sphere, but lags behind it by an angle $\pi/4m$. The variation of the angle through which the field lags as a function of $4\pi k \omega b^2$ is shown in figure 5 for $n = 2$, $m = 2$.

In Elsasser's notation this process is the interaction of a T_n^{mc} field with a T_1 motion to produce a T_n^{ms} field. If the lag is $\pi/4m$, an equal mixture of the two fields is produced. From figure 5 it is found that T_2^{2s} exceeds 50 % of T_2^{2c} for all values of $4\pi k \omega b^2$ exceeding 9.2. This argument suggested the statement made in § 8 of II that this interaction is strong if $4\pi k \omega b^2 > 1$, though the analogy with the system there considered is not close enough for more than a qualitative comparison.

The field produced by a stationary source at the centre of a rotating sphere has also been considered. For fast rotation the field is powerfully screened; at $r = b$ it

contains a term $\exp[-\sqrt{(2\pi\kappa\omega mb^2)}]$. The field at $r = b$ is rotated through an angle $\sqrt{(2\pi\kappa\omega mb^2)} + \frac{1}{4}n\pi$. This angle increases indefinitely with ω in contrast to the result obtained when the source rotates with the sphere.

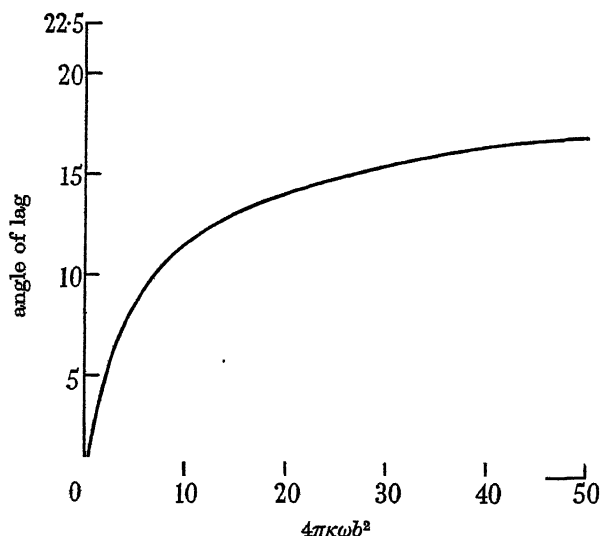


FIGURE 5. Coupling of T_2^{2c} and T_2^{2s} .

Source of current at infinity

The external field due to a current source at a great distance can be represented by terms such as $C\bar{H}_{Tnm}$. The boundary conditions can be satisfied by adding a field $C'\bar{H}'_{Tnm}$ outside the sphere and a field DH_{Tnm} inside. It is found that the variation of the magnitude and direction of the field at $r = b$ with velocity of rotation is controlled by a factor $J_{n+\frac{1}{2}}(kb)/kb J_{n+\frac{1}{2}}(kb)$ whose modulus and phase give the intensity and direction of the field. For fast rotation the whole field at $r = b$ is rotated through an angle $\frac{1}{2}\pi/m$ relative to that at a great distance or to that obtained without rotation.

9. ELECTROMAGNETIC COUPLES

In the systems discussed in §§ 7 and 8 electric currents are maintained by electromagnetic induction in a rotating sphere. The energy of these currents must be supplied by external forces driving the sphere. The interaction of the current and the field must therefore produce a couple tending to stop the sphere. This couple, Γ , is given by

$$\Gamma = \iiint r^3 \sin^2 \theta (I_\theta H_r - I_r H_\theta) d\theta d\phi dr, \quad (28)$$

where the integral extends through the rotating sphere. Alternatively, it may be calculated directly from the dissipation of energy,

$$\omega\Gamma = \frac{1}{\kappa} \iiint I^2 r^2 \sin \theta d\theta d\phi dr,$$

where the integral extends through the sphere and the shell. The two expressions for Γ may easily be shown to be equivalent, but (28) is the more convenient, as it has only to be taken through the sphere.

The simplest problem, and that which is most important for the applications of II, is to calculate the couple when a sphere is placed in a uniform field parallel to its axis of rotation. The toroidal field discussed in § 6 is then in the ϕ direction and can produce no couple about the axis of rotation. The couple is produced entirely by the interaction of the uniform field F and the meridional current (26). The integral (28) is then easily evaluated and gives

$$\Gamma = \frac{8\pi}{75} F^2 b^5 \kappa \omega (1 - b^5/a^5) = \frac{4}{15} F H_\phi (\text{max.}) b^3. \quad (29)$$

The couple thus increases indefinitely with ω . The rate at which work must be done on the sphere to keep its angular velocity constant is $\Gamma\omega$. If $a \gg b$ the whole energy in the toroidal field is shown in § 6 to be $8b^2 H_\phi (\text{max.})/63$. The whole energy of the field is thus replaced in a time $30H_\phi (\text{max.})/63F\omega = \pi^2 \tau_0/21$, where $\tau_0 = 4\kappa b^2/\pi$ is the time of relaxation of the gravest mode of free decay of a sphere of radius b (§ 11). Since the initial rate of decay of energy in the free mode W/W is $2/\tau_0$, the fraction $\pi^2/21$ would be expected to be near a half. Actually it is 0.47; exact agreement is not to be expected, as the radial variation of H_ϕ is not the same for the steady and the freely decaying fields.

Since the moment of inertia of the sphere is $8\pi\Delta b^2/15$, where Δ is its density, (29) may be written

$$\Gamma = \frac{1}{5} F^2 I \kappa \omega (1 - b^5/a^5)/\Delta.$$

The angular deceleration that occurs if the maintaining couple is removed is therefore independent of the scale of the system. This is in marked contrast to the couple on a sphere rotating in a viscous fluid where the couple is only proportional to b^3 and the deceleration to $1/b^2$. On a laboratory scale the effects of viscosity on the motion of a fluid much exceed those of the electromagnetic forces in fields of a few gauss. In large bodies, such as the core of the earth or the stars, this is no longer true, and motions such as that due to convection will be controlled by electromagnetic forces rather than by viscosity. The resulting magneto-hydrodynamics offers a promising field for investigation (cf. Alfvén 1947 and earlier papers).

If the applied external field is $F\bar{H}_{Snm}$, the field and current in the sphere are given by (24). Substituting in (28) and omitting terms in ${}_r I_{Tnm}$ and ${}_r H_{T(n\pm 1)m}$, which are identically zero, gives

$$\Gamma = \iiint r^3 \sin^2 \theta [(A_\theta I_{Tnm} + D_{n+1\theta} I_{S(n+1)m} + D_{n-1\theta} I_{S(n-1)m}) A_r H_{Snm} - (D_{n+1r} I_{S(n+1)m} + D_{n-1r} I_{S(n-1)m}) (A_\theta H_{Snm} + D_{n+1\theta} H_{T(n+1)m} + D_{n-1\theta} H_{T(n-1)m})] dr d\theta d\phi. \quad (30)$$

The fields used in this integral must be real fields. It is not allowable to use the complex A , D , H and I 's and to take the real part of Γ at the end. A term AH must therefore be taken as $\frac{1}{2}(AH + A^*H^*)$, where A^* and H^* are the conjugate complexes to A and H . These are obtained by putting $-i$ for i and $k^* = ik$ for k in the expressions of table 1 and § 6.

The terms containing the product of any two D 's vanish on integration, leaving only those in two A 's or an A and a D . That is, the couples are due to the interaction of the field AH_{Snm} with the current AI_{Tnm} or with one of the currents $D_{n\pm 1}I_{S(n\pm 1)m}$ producing the toroidal field. These couples will be called Γ_1 , Γ_2 and Γ_3 . The θ and ϕ integrations involve integrals of the form $\iint P_{n\pm 1}^m(\theta) \frac{dP_n^m(\theta)}{d\theta} \sin \theta d\theta d\phi$, which are easily evaluated. The r integrations involve products of Bessel functions, all of which can be reduced to integrals such as

$$2k^2 \int_0^b r |J_{n+\frac{1}{2}}(kr)|^2 dr = b[k^* J_{n+\frac{1}{2}}(kb) J_{n-\frac{1}{2}}(k^*b) - k J_{n+\frac{1}{2}}(k^*b) J_{n-\frac{1}{2}}(kb)].$$

Γ_1 is independent of a , but Γ_2 and Γ_3 vanish if $b = a$ and there is no shell. The expressions for finite a are complicated and we give only those for $a \gg b$. If $m \neq 0$ these are found, after lengthy reductions, to be

$$\left. \begin{aligned} \Gamma_1 &= -\frac{m^{\frac{1}{2}} n(2n+1)(n+m)! F^2 b^2 M_{n+\frac{1}{2}}}{4(n+1)(n-m)! \sqrt{(\pi k \omega)} M_{n-\frac{1}{2}}} \cos(\zeta_{n+\frac{1}{2}} - \zeta_{n-\frac{1}{2}} - \tfrac{1}{4}\pi), \\ \Gamma_2 &= -\frac{n^2(n+2)(n-m+1)(n+m+1)! F^2 b^2 M_{n+\frac{1}{2}}}{4(n+1)(2n+3)(n-m)! m^{\frac{3}{2}} \sqrt{(\pi k \omega)} M_{n-\frac{1}{2}}} \cos(\zeta_{n+\frac{1}{2}} - \zeta_{n-\frac{1}{2}} - \tfrac{1}{4}\pi), \\ \Gamma_3 &= \frac{n(n-1)(n+m)! (n^2-m^2) F^2 b^2 M_{n+\frac{1}{2}}}{4(2n-1)(n-m)! m^{\frac{3}{2}} \sqrt{(\pi k \omega)} M_{n-\frac{1}{2}}} \left[\sin(\zeta_{n+\frac{1}{2}} - \zeta_{n-\frac{1}{2}}) \cos(\zeta_{n-\frac{1}{2}} - \zeta_{n-\frac{1}{2}} + \tfrac{1}{4}\pi) \right. \\ &\quad \left. + \frac{M_{n+\frac{1}{2}}}{M_{n-\frac{1}{2}}} \cos(\zeta_{n+\frac{1}{2}} - \zeta_{n-\frac{1}{2}}) \cos(\zeta_{n+\frac{1}{2}} - \zeta_{n-\frac{1}{2}} - \tfrac{1}{4}\pi) \right], \end{aligned} \right\} \quad (31)$$

where $M_{n+\frac{1}{2}}$ and $\zeta_{n+\frac{1}{2}}$ are the modulus and phase of $J_{n+\frac{1}{2}}(kb)$. For small ω the first two terms of the ascending series may be used for the Bessel functions. All these expressions then simplify and become

$$\left. \begin{aligned} \Gamma_1 &= \frac{2m^2 n(n+m)! b^5 F^2 \pi k \omega}{(2n+1)(2n+3)(n+1)(n-m)!}, \\ \Gamma_2 &= \frac{2n^2(n+2)(n-m+1)(n+m+1)! b^5 F^2 \pi k \omega}{(n+1)(2n+1)^2(2n+3)^2(n-m)!}, \\ \Gamma_3 &= \frac{2(n-1)n(n+m)! (n^2-m^2) b^5 F^2 \pi k \omega}{(2n-1)^2(2n+1)^2(n-m)!}. \end{aligned} \right\} \quad (32)$$

As ω increases the couples increase, pass through a maximum and finally decrease. For large values of $k\omega b^2$ they become

$$\begin{aligned} \Gamma_1 &= \frac{m^{\frac{1}{2}} n(2n+1)(n+m)! b^2 F^2}{4(n+1)(n-m)! \sqrt{(2\pi k \omega)}}, \\ \Gamma_2 &= \frac{n^2(n+2)(n-m+1)(n+m+1)! b^2 F^2}{4(n+1)(2n+3)(n-m)! m^{\frac{3}{2}} \sqrt{(2\pi k \omega)}}, \\ \Gamma_3 &= \frac{(n-1)n(n+m)! (n^2-m^2) b^2 F^2}{4(2n-1)(n-m)! m^{\frac{3}{2}} \sqrt{(2\pi k \omega)}}. \end{aligned}$$

The sphere rotating in a uniform field at right angles to its axis ($n = m = 1$) has been investigated numerically using (31). The results are shown in figure 6. For slow rotation

$$\Gamma_1 = 2\pi b^5 F^2 \kappa \omega / 15, \quad \Gamma_2 = 2\pi b^5 F^2 \kappa \omega / 25, \quad \Gamma_3 = 0.$$

This result for Γ_1 may be obtained immediately by considering the couple produced by the field F on the dipole $2\pi b^2 F \kappa \omega / 15$ of § 5. For fast rotation

$$\Gamma_1 = 3b^2 F^2 / 4 \sqrt{(2\pi \kappa \omega)}, \quad \Gamma_2 = 9b^2 F^2 / 20 \sqrt{(2\pi \kappa \omega)}, \quad \Gamma_3 = 0.$$

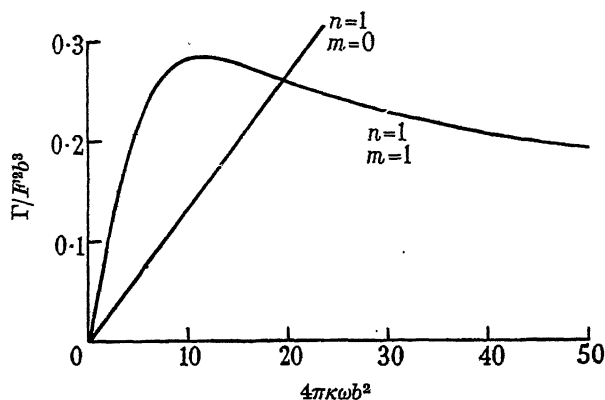


FIGURE 6. Couple on a sphere rotating in a uniform field.

The decrease of the couple with increasing speed of rotation is at first sight unexpected, but may easily be explained. Since for large ω the induced moment is independent of ω and is produced by currents almost entirely confined to a layer of thickness proportional to $1/\sqrt{(\kappa\omega)}$, the current density is proportional to $\sqrt{(\kappa\omega)}$. The rate of dissipation of energy per unit volume is thus independent of κ and is proportional to ω . The total rate of dissipation of energy is proportional to $\sqrt{(\omega/\kappa)}$ and the couple to $1/\sqrt{(\omega\kappa)}$. The decrease of the couple with increasing speed of rotation is thus accounted for. Alternatively, the decrease in Γ_1 may be regarded as due to the moment tending to move into a direction antiparallel to the inducing field, which can therefore, in the limit, exert no couple on it.

The couples (for any ω) for $m = 0$ are twice those obtained by putting $m = 0$ in (32). The difference is caused by certain terms that are present when $m = 0$, vanishing on integration when $m \neq 0$. Γ_1 is always zero when $m = 0$ owing to the vanishing of the induced multipole. Γ_2 and Γ_3 are given by

$$\Gamma_2 = \frac{4\pi n^2(n+1)(n+2)F^2 b^5 \kappa \omega}{(2n+1)^2(2n+3)^2},$$

$$\Gamma_3 = \frac{4\pi n^3(n-1)F^2 b^5 \kappa \omega}{(2n+1)^2(2n-1)^2}.$$

The result for $n = 1$ is shown in figure 6; it agrees with that obtained independently above and provides a check on the elaborate calculation. The result for $n = 2, m = 1$ has also been calculated independently as a further check.

This completes the discussion of the couples produced by an external field $F\bar{H}_{Snm}$. The main result is that couples large compared to F^2b^3 can be produced only by the part of the external field that is symmetrical about the axis ($m = 0$). The couple due to this increases proportionally to $F\kappa\omega b^5$.

If a field $F_1\bar{H}_{S\nu\mu}$ is present as well as $F\bar{H}_{Snm}$, an examination of (30) shows that the couple produced is not always the sum of those produced by the two fields separately. Additional terms can arise only if $m = \mu$ and $\nu = n \pm 2$. They have not been considered in detail.

The couples produced by a stationary magnetic multipole at the centre of the sphere can also be calculated from the results of § 7. Only the case $m = 0$ has been considered in detail. The couple is then caused entirely by the interaction of the multipole with the current that produces the toroidal field. The result is

$$\Gamma = \frac{4(n+1)(n+2)b^5F_1^2\pi\kappa\omega}{(2n+1)^2(2n+3)^2},$$

where F_1 is the field produced by the multipole at $\theta = 0$, $r = b$. For $n = 1$

$$\Gamma = 8\pi b^5 F_1^2 \kappa \omega.$$

This is identical with the previous result (29) for a uniform field; the change from a uniform field to a dipole with a singularity at the origin has produced no change in couple. From this it appears that the couple is insensitive to the radial distribution of the inducing field. This is important for the applications of II, where the results of induction by a uniform field are applied to problems in which the field has the same angular symmetry, but an unknown radial variation.

10. INDUCTION IN AN OSCILLATING SPHERE

There is nothing in the derivation of (6) that requires the angular velocity, ω , to be constant. If it is a function of the time the variables will still separate to give (8) provided (7) is replaced by

$$H = H_1(r, \theta) e^{im\phi} T(t),$$

where

$$\dot{T} + (k^2/4\pi\kappa + i\omega m) T = 0,$$

which gives

$$T = T_0 \exp\left(-k^2 t/4\pi\kappa - im \int \omega dt\right). \quad (33)$$

In these expressions k is an adjustable constant, but is no longer given by (9). The expression (33) is well adapted to the study of the free decay of the field in an oscillating or accelerating sphere surrounded by a conducting shell. All that is necessary is to choose k so as to satisfy the boundary conditions as in § 11 below. The solutions corresponding to a steady field at infinity cannot easily be obtained from it unless $m = 0$. If $m = 0$, ω does not occur in (33). It does, however, occur in the boundary conditions for the electric field.

For simplicity suppose $n = 1$, $m = 0$ and that $\omega = \omega_0 e^{i\omega t}$, by analogy with § 6 it is natural to try to obtain a solution in which the magnetic field outside the shell is

$F\bar{H}_{S1} + B\bar{H}'_{S1}$, that in the shell $F\bar{H}'_{S1} + B\bar{H}_{S1} + CH_{T2} + C'H'_{T2}$, and that in the sphere $AH_{S1} + DH_{T2}$. These terms may not all have the same value of the constant k . If $F\bar{H}_{S1}$ is constant in time the boundary conditions for H will require the H_{S1} and $\bar{H}_{S1}$ terms to be constants also and σ , and thus k , to be zero for these terms. In the expression from table 1 for E_θ , A is multiplied by ω and is therefore proportional to $e^{i\omega t}$. If the boundary conditions are to be satisfied the terms in C and D must also contain a factor $e^{i\omega t}$. From (33) this implies that $k^2 = -4\pi i\nu\kappa$ for these terms. These considerations lead to the following fields and currents:

$$\left. \begin{aligned} H_r &= A \cos \theta, \\ H_\theta &= -A \sin \theta, \\ H_\phi &= 3D e^{i\omega t} (kr)^{-\frac{1}{2}} J_{\frac{3}{2}}(kr) \sin \theta \cos \theta, \\ I_r &= (3/4\pi) D e^{i\omega t} (kr)^{-\frac{1}{2}} J_{\frac{3}{2}}(kr) (3 \cos^2 \theta - 1), \\ E_\theta &= -[A\omega_0 e^{i\omega t} r + (3/4\pi\kappa) D e^{i\omega t} (kr)^{-\frac{1}{2}} \{kr J_{\frac{3}{2}}(kr) - 2J_{\frac{3}{2}}(kr)\} \cos \theta \sin \theta], \\ H_r &= (F - 2Bb^3/r^3) \cos \theta, \\ H_\theta &= -(F + Bb^3/r^3) \sin \theta, \\ H_\phi &= 3(kr)^{-\frac{1}{2}} e^{i\omega t} [CJ_{\frac{3}{2}}(kr) + C'Y_{\frac{3}{2}}(kr)] \cos \theta \sin \theta, \\ I_r &= (3/4\pi) (kr)^{-\frac{1}{2}} e^{i\omega t} [CJ_{\frac{3}{2}}(kr) + C'Y_{\frac{3}{2}}(kr)] (3 \cos^2 \theta - 1), \\ E_\theta &= -(3/4\pi\kappa) e^{i\omega t} (kr)^{-\frac{1}{2}} [C\{(kr) J_{\frac{3}{2}}(kr) - 2J_{\frac{3}{2}}(kr)\} + C'\{kr Y_{\frac{3}{2}}(kr) - 2Y_{\frac{3}{2}}(kr)\}] \\ &\quad \times \cos \theta \sin \theta, \end{aligned} \right\} \begin{array}{l} \text{inside the} \\ \text{sphere,} \end{array}$$

$$\left. \begin{aligned} H_r &= (F - 2Bb^3/r^3) \cos \theta, \\ H_\theta &= -(F + Bb^3/r^3) \sin \theta, \\ H_\phi &= 0, \\ E_\theta &= 3G e^{i\omega t} (b^3/r^2) \sin \theta \cos \theta, \end{aligned} \right\} \text{outside.}$$

The boundary conditions are all satisfied if

$$A = F, \quad B = 0, \quad C = -C'Y_{\frac{3}{2}}(ka)/J_{\frac{3}{2}}(ka), \quad D = C'[Y_{\frac{3}{2}}(kb)/J_{\frac{3}{2}}(kb) - Y_{\frac{3}{2}}(ka)/J_{\frac{3}{2}}(ka)],$$

$$C' = \frac{2}{3}\pi^2 F\omega\kappa b^2 \sqrt{(kb)} J_{\frac{3}{2}}(kb).$$

These expressions enable all the fields to be calculated in terms of F . The maximum toroidal field occurs at $r = b$, $\theta = \frac{1}{2}\pi$; its amplitude H_ϕ (max.) and phase ζ are given by

$$H_\phi (\text{max.}) e^{i\zeta} = \pi^2 F\omega_0 b^2 \kappa J_{\frac{3}{2}}^2(kb) [Y_{\frac{3}{2}}(kb)/J_{\frac{3}{2}}(kb) - Y_{\frac{3}{2}}(ka)/J_{\frac{3}{2}}(ka)]. \quad (34)$$

If the period of oscillation of the sphere, τ , is long compared to the time of relaxation of the shell $4\kappa a^2/\pi$, then ka and kb will be small and the Bessel functions in (34) may be replaced by the first term of their expansions in ascending powers of ka and kb ; the result is

$$H_\phi (\text{max.}) = -\frac{2}{3}\pi F\omega_0 b^2 \kappa (1 - b^5/a^5),$$

which is identical with the result (25b) obtained for a steadily rotating sphere. This agreement was to be expected, since a sphere oscillating with a very long period simulates a one rotating steadily.

If the oscillation is rapid compared to the time of relaxation, kb and ka will be large and the Bessel functions may be replaced by the first term of the terminating series. If $b \gg a$ the result when $\omega = \omega_0 \cos 2\pi t/\tau$ is

$$H_\phi = F\omega_0 b^2 \kappa \sqrt{(\tau/2\pi\tau_0)} \cos(2\pi t/\tau - \frac{1}{4}\pi),$$

where $\tau_0 = 4\kappa b^2/\pi$ is the greatest time of relaxation of a stationary sphere of radius b . The amplitude is $(5/2^{\frac{1}{2}}\pi^{\frac{1}{2}}) \sqrt{(\tau/\tau_0)} = 0.32 \sqrt{(\tau/\tau_0)}$ times the result for steady rotation with angular velocity ω_0 . The physical reason for the reduction when $\tau \gg \tau_0$ is that the current system has not time to establish itself before the rotation is reversed and the field has to decrease and build up in the opposite direction.

The toroidal field inside the sphere varies as $J_{\frac{3}{2}}(kr)/\sqrt{(kr)}$. For large kr the modulus of this is equal to $\exp\{2\pi r \sqrt{(\kappa/\tau)}\}/8\pi r \sqrt{(2\kappa/\tau)}$. Thus the field falls to $1/e$ of its value at $r = b$ at a depth approximately equal to $\sqrt{(\tau/4\pi^2\kappa)} = b \sqrt{(\tau/\pi^2\tau_0)}$. If the shell is thin compared to this the reduction of field due to rapid oscillation does not occur and $H_\phi(\text{max.})$ is found to be

$$H_\phi(\text{max.}) = -2\pi\kappa a^2 \omega_0 H_0(1 - b/a) \cos 2\pi t/\tau.$$

The amplitude of this is identical with that for steady rotation when $(a - b) \gg b$.

To illustrate the behaviour of (34) over the whole range of values of τ/τ_0 , $H_\phi(\text{max.})$ has been calculated for $a = 2b$. The results are illustrated in figure 7, where the ratio of the toroidal field to that for steady rotation is plotted against τ/τ_0 .

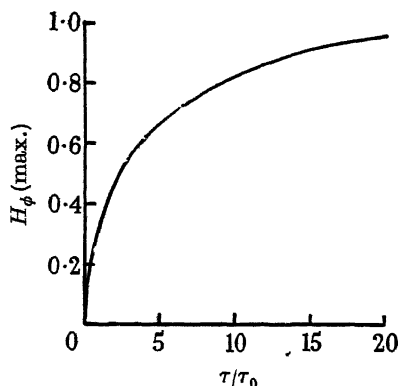


FIGURE 7. Toroidal field in an oscillating sphere. $H_\phi(\text{max.})$ is the maximum value of the field expressed as a fraction of the value it would have for a continuous rotation with a velocity equal to the maximum occurring in the oscillation.

The results of this section are probably typical of what would be obtained for any oscillatory motion. In particular, they suggest that the magnetic effects of tidal motion in the earth's core will be small, since the ratio τ/τ_0 will be that of a day to 14,000 years. This result is used in II as a reason for rejecting a theory of the earth's magnetic field depending on tidal motions in the interior of the core.

11. THE FREE MODES

The problems treated in §§5 to 10 have all involved a specified inducing field maintained by a singularity of field or current either at the origin or at infinity. We now discuss the freely decaying fields that occur when there is no externally main-

tained field. These may be resolved into normal modes, for each of which the field decays so that its ratio at any two points is independent of the time. Any initial field may be built up by a combination of these normal modes. The modes belong to two series, one in which the magnetic field is toroidal and a second in which both poloidal and toroidal fields are present. The treatment of the latter is considerably complicated by a coincidence of its periods with those of the first series. We therefore first discuss the modes with a purely toroidal magnetic field.

Toroidal magnetic field

Suppose the magnetic field inside the sphere to be AH_{Tnm} . There can be no corresponding toroidal field in the insulator outside the shell. If the shell is of zero thickness the boundary conditions for H require H_{Tnm} to vanish at $r = a = b$, and thus from table 1

$$J_{n+\frac{1}{2}}(ka) = 0. \quad (35)$$

The boundary conditions for the electric field can be satisfied by an external electric field $\bar{E}'_{Snm}$.

The solutions of (35) are all real. Call them k_s , where $s = 1, 2, \dots$. The field decays as $e^{-\sigma t}$, where from (9)

$$\sigma_s = k_s^2/4\pi\kappa + i\omega m.$$

The magnetic field is given by

$$\begin{aligned} H_r &= 0, \\ H_\theta &= iAm \exp[-k_s^2 t/4\pi\kappa] (k_s r)^{-\frac{1}{2}} J_{n+\frac{1}{2}}(k_s r) \exp[i m(\phi - \omega t)] P_n^m(\theta)/\sin \theta, \\ H_\phi &= -A \exp[-k_s^2 t/4\pi\kappa] (k_s r)^{-\frac{1}{2}} J_{n+\frac{1}{2}}(k_s r) \exp[i m(\phi - \omega t)] dP_n^m(\theta)/d\theta. \end{aligned}$$

The real part of this is

$$\begin{aligned} H_\theta &= -Am \exp[-k_s^2 t/4\pi\kappa] \sin m(\phi - \omega t) (k_s r)^{-\frac{1}{2}} J_{n+\frac{1}{2}}(k_s r) P_n^m(\theta)/\sin \theta, \\ H_\phi &= -A \exp[-k_s^2 t/4\pi\kappa] \cos m(\phi - \omega t) (k_s r)^{-\frac{1}{2}} J_{n+\frac{1}{2}}(k_s r) dP_n^m(\theta)/d\theta. \end{aligned}$$

The field therefore decays with a time constant $4\pi\kappa/k^2$ and rotates with the sphere. The time of relaxation is, as was to be expected, the same as for a stationary sphere.

If the sphere is surrounded by a concentric conducting shell of radius a , the result is considerably more complicated. If the magnetic field in the sphere is AH_{Tnm} , that in the shell will be $BH_{Tnm} + B'H'_{Tnm}$. Outside the shell the magnetic field is zero and there is an electric field $G\bar{E}'_{Snm}$; the time of relaxation, $1/\sigma$, must be the same in the sphere and in the shell. This requires

$$\sigma = k^2/4\pi\kappa + i\omega m = k_1^2/4\pi\kappa, \quad (36)$$

where k and k_1 are the values of k in the sphere and shell. All the boundary conditions are satisfied if

$$\begin{aligned} B' &= -BJ_{n+\frac{1}{2}}(k_1 a)/Y_{n+\frac{1}{2}}(k_1 a), \\ \frac{A}{B} &= \left(\frac{k}{k_1}\right)^{\frac{1}{2}} \left[\frac{J_{n+\frac{1}{2}}(k_1 b)}{Y_{n+\frac{1}{2}}(k_1 b)} - \frac{J_{n+\frac{1}{2}}(k_1 a)}{Y_{n+\frac{1}{2}}(k_1 a)} \right] \frac{Y_{n+\frac{1}{2}}(k_1 b)}{J_{n+\frac{1}{2}}(kb)} = \left(\frac{k_1}{k}\right)^{\frac{1}{2}} \left[\frac{J_{n-\frac{1}{2}}(k_1 b)}{Y_{n-\frac{1}{2}}(k_1 b)} - \frac{J_{n+\frac{1}{2}}(k_1 a)}{Y_{n+\frac{1}{2}}(k_1 a)} \right] \frac{Y_{n-\frac{1}{2}}(k_1 b)}{J_{n-\frac{1}{2}}(kb)}. \end{aligned} \quad (37)$$

The second half of (37) combined with (36) fixes a series of possible values of k , k_1 and σ . The other two relations then give B and B' in terms of A . The A 's corresponding to each value of k will depend on the distribution of field at $t = 0$.

If $m = 0$, $k = k_1$ and (37) becomes

$$\frac{Y_{n+\frac{1}{2}}(kb) J_{n-\frac{1}{2}}(kb) - J_{n+\frac{1}{2}}(kb) Y_{n-\frac{1}{2}}(kb)}{J_{n+\frac{1}{2}}(kb) J_{n-\frac{1}{2}}(kb)} \frac{J_{n+\frac{1}{2}}(ka)}{Y_{n+\frac{1}{2}}(ka)} = 0,$$

which implies either

$$Y_{n+\frac{1}{2}}(kb) J_{n-\frac{1}{2}}(kb) - J_{n+\frac{1}{2}}(kb) Y_{n-\frac{1}{2}}(kb) = 0$$

or

$$J_{n+\frac{1}{2}}(ka) = 0.$$

The first has no finite solution since it is equal to $-2/(\pi kb)$. The second is identical with the result for a stationary sphere of radius a . The decay of a toroidal field symmetrical about the axis is therefore unaffected by rotation. B' is zero and the field is AH_{Tn} throughout both sphere and shell. The discontinuity in the motion at $r = b$ has no effect whatever on the field. This is to be expected as the magnetic field is entirely in the direction of the motion.

When $m \neq 0$ (36) shows that at least one of k and k_1 must be complex; (37) will then ensure that, except in special cases, both are complex. The ϕ component of the field in the sphere will be

$$A(kr)^{-\frac{1}{2}} J_{n+\frac{1}{2}}(kr) \exp[-(k^2/4\pi\kappa + i\omega m)t] e^{im\phi} dP_n^m(\theta)/d\theta.$$

If the real and imaginary parts of $k^2/4\pi\kappa$ are α and β , the variation with t and ϕ will give a term $e^{-\alpha t} \exp\{im[\phi - (\omega + \beta/m)t]\}$. The angular velocity with which the field rotates is therefore $\omega + \beta/m$ and differs from that of the sphere. For a given n and m the angular velocity is necessarily the same at all points both in the sphere and in the shell.

The relation (37), which with (36) determines k and k_1 , may be written

$$\frac{J_{n+\frac{1}{2}}(k_1 a)}{Y_{n+\frac{1}{2}}(k_1 a)} = \frac{k J_{n+\frac{1}{2}}(k_1 b)/J_{n+\frac{1}{2}}(kb) - k_1 J_{n-\frac{1}{2}}(k_1 b)/J_{n-\frac{1}{2}}(kb)}{k Y_{n+\frac{1}{2}}(k_1 b)/Y_{n+\frac{1}{2}}(kb) - k_1 Y_{n-\frac{1}{2}}(k_1 b)/Y_{n-\frac{1}{2}}(kb)}. \quad (38)$$

A discussion of (38) is difficult in the absence of tables of $J_{n+\frac{1}{2}}$ for general complex argument. The behaviour of its roots may be seen from a consideration of simple cases. If the sphere is stationary, $k = k_1$ and the right-hand side is zero. To make the left-hand side also zero $J_{n+\frac{1}{2}}(k_1 a) = J_{n+\frac{1}{2}}(ka) = 0$ as in (35). For slow rotation the difference between k and a root of (35) will be a small quantity, and the Bessel functions may be expanded in a Taylor series involving only functions of real argument. From this it is found that, to the first order of small quantities, the time of relaxation is the same as for a stationary sphere, but that the field rotates with an angular velocity Ω given by

$$\Omega/\omega = (\pi qb/2a)^2 Y_{n+\frac{1}{2}}^2(q) [J_n^2(qb/a) - J_{n+\frac{1}{2}}(q/ba) J_{n-\frac{1}{2}}(qb/a)], \quad (39)$$

where q is a root of $J_{n+\frac{1}{2}}(q) = 0$; by a known expansion (Watson 1944, p. 152) this becomes

$$\Omega/\omega = \pi^2 Y_{n+\frac{1}{2}}^2(q) \sum_{l=0}^{\infty} (n+2m+\frac{3}{2}) J_{n+\frac{1}{2}+2l}^2(qb/a). \quad (40)$$

This is necessarily positive, and the field therefore rotates in the same direction as the sphere. If $b = 0$ (39) gives $\Omega = 0$, if $b = a$ an application of the recurrence relations gives $\Omega = \omega$. The behaviour of the function between these extremes may be seen by differentiating (40) with respect to b . After some rearrangement the result becomes

$$\frac{d(\Omega/\omega)}{d(b/a)} = \frac{1}{2}\pi^2(q^2b/a) Y_{n+\frac{1}{2}}^2(q) J_{n+\frac{1}{2}}^2(qb/a).$$

This is necessarily positive, and is zero when $b = 0$ or a . The velocity with which the field rotates therefore increases continuously as b increases from zero to a . Figure 8 shows the variation for the lowest mode with $n = 1$ ($q = 4.4934$). It seems likely on physical grounds that a similar behaviour will occur when ω is not small, but I have not been able to prove this from (36) and (37).

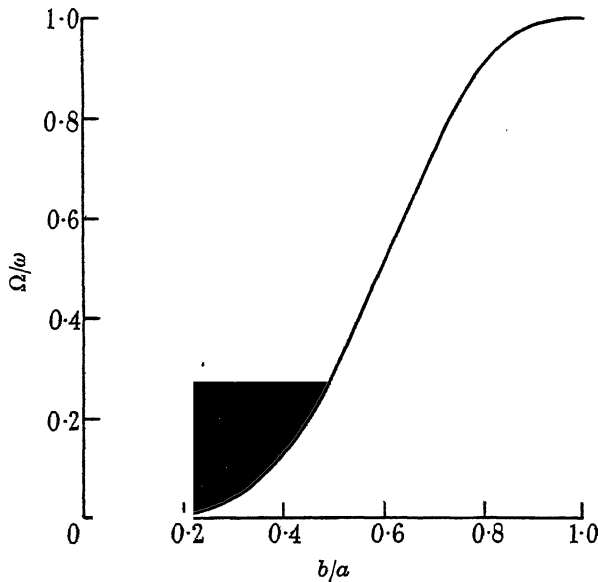


FIGURE 8. Rate of rotation of field in a free mode with $n = 1$.

It is conceivable that (38) might possess solutions other than those corresponding to the roots of (35). In particular, it is desirable to investigate the possibility of roots for which kb remains finite when $a \rightarrow \infty$. These would be modes of a rotating sphere in an indefinitely extended conducting medium. We now show that such modes do not exist when $n = 1$.

When $n = 1$ (38) becomes

$$\frac{\tan k_1 a - k_1 a}{k_1 a \tan k_1 a + 1} = \frac{k/k_1^2 b - 1/kb - (k/k_1) \cot k_1 b + \cot kb}{k/k_1^2 b - 1/kb + (k/k_1) \tan k_1 b + \cot kb} \tan k_1 b. \quad (41)$$

The left-hand side only tends to a limit with increasing a if k_1 is complex or imaginary. If the imaginary part of k_1 is positive, the limit is i ; if it is negative it is $-i$. (42) then factorizes and becomes

$$(\tan k_1 b \mp i) (\pm ik/k_1 - k/k_1^2 b + 1/kb - \cot kb) = 0.$$

The first factor has no finite root. If k_1 be eliminated from the second factor by the use of (36) it gives

$$\cot kb = \pm kb/(k^2b^2 + i\rho)^{\frac{1}{2}} + i\rho/kb(k^2b^2 + i\rho),$$

where $\rho = 4\pi\kappa\omega b^2$. A numerical study of the equation suggests that it has no roots other than the irrelevant one $k = 0$. It therefore appears that there are no modes with finite k when a is very large. This is likely to be true for all n but has not been proved.

Poloidal magnetic field

If the magnetic field in the sphere is AH_{Snm} , the boundary conditions for H and I can be satisfied by a field $BH_{Snm} + B'H'_{Snm}$ in the shell and $F\bar{H}'_{Snm}$ outside. These boundary conditions give B , B' and F in terms of A and a condition that must be satisfied by k and k_1 in the sphere and shell. The latter relation is closely similar to that found for the toroidal field (38); it is

$$\frac{J_{n-\frac{1}{2}}(k_1 a)}{Y_{n-\frac{1}{2}}(k_1 a)} = \frac{k J_{n-\frac{1}{2}}(k_1 b) J_{n+\frac{1}{2}}(kb) - k_1 J_{n+\frac{1}{2}}(k_1 b) J_{n-\frac{1}{2}}(kb)}{k Y_{n-\frac{1}{2}}(k_1 b) J_{n+\frac{1}{2}}(kb) - k_1 Y_{n+\frac{1}{2}}(k_1 b) J_{n-\frac{1}{2}}(kb)}, \quad (42)$$

if $b = a$ or $m = 0$ this reduces to

$$J_{n-\frac{1}{2}}(ka) = 0, \quad (43)$$

which is also the result for a stationary sphere.

Just as in §6 the above fields will not satisfy the boundary conditions for the electric field, and it is necessary to introduce toroidal magnetic fields $D_{n\pm 1}H_{T(n\pm 1)m}$ inside the sphere, $C_{n\pm 1}H_{T(n\pm 1)m} + C'_{n\pm 1}H'_{T(n\pm 1)m}$ in the shell and $E_{n\pm 1}\bar{H}'_{T(n\pm 1)m}$ outside. The electric fields corresponding to these enable the boundary conditions for E to be satisfied without disturbing (42) or the values of A , B , B' or F . The presence of these terms means that the poloidal magnetic field is not a normal mode of the system; it is necessarily accompanied by a toroidal field. By analysis closely similar to that of §6 the values of the C 's, D 's and E 's can be found in terms of A . If $m \neq 0$ the results are complicated, but present no difficulty of principle. If $m = 0$ the coefficient of C_{n-1} vanishes in the relation connecting it with A and the method breaks down. This can be traced to the value of k given by (43) also satisfying (35) with a value of n reduced by one unit. That is to say, the k belonging to the toroidal mode of order $(n-1)$ is equal to that for the mode of order n in the poloidal series. Similar identities among the 'periods' of normal modes are common in dynamical systems, and the analogy suggests a way out of the difficulty. What is required is a solution corresponding to the $t e^{-kt}$ solution of $\ddot{x} + 2k\dot{x} + k^2x = 0$. This solution may be obtained by differentiating the simpler solution e^{-kt} with respect to k . This suggests differentiating the solutions of tables 1 and 2 with respect to k remembering that, since $m = 0$, $\sigma = k^2/4\pi\kappa$. Differentiating ${}_pH_{Tn}$ gives

$$\frac{dH_\phi}{dk} = \left[\frac{k}{2\pi\kappa} \{ t J_{n+\frac{1}{2}}(kr)/(kr)^{\frac{1}{2}} + \frac{1}{2}(kr)^{\frac{1}{2}} J_{n-\frac{1}{2}}(kr)/2\sigma \} + (n+1) J_{n+\frac{1}{2}}(kr)/k^{\frac{1}{2}} r^{\frac{1}{2}} \right] e^{-\sigma t} \frac{dP_n(\theta)}{d\theta}.$$

Since the part in $J_{n+\frac{1}{2}}(kr)/r^{\frac{1}{2}}$ is already known to be a solution this suggests that the new solution may conveniently be taken as a constant times the term in $\{ \}$. Substituted in (1) this gives

$$\left. \begin{aligned} H_r &= 0, \\ H_\theta &= 0, \\ H_\phi &= [tJ_{n+\frac{1}{2}}(kr)/(kr)^{\frac{1}{2}} - (kr)^{\frac{1}{2}}J_{n-\frac{1}{2}}(kr)/2\sigma] e^{-\sigma t} dP_n(\theta)/d\theta, \\ \kappa E_r = I_r &= -n(n+1)(4\pi r)^{-1} [tJ_{n+\frac{1}{2}}(kr)/(kr)^{\frac{1}{2}} - (kr)^{\frac{1}{2}}J_{n-\frac{1}{2}}(kr)] e^{-\sigma t} P_n(\theta), \\ \kappa E_\theta = I_\theta &= (4\pi r)^{-1} \{ nJ_{n+\frac{1}{2}}(kr)/(kr)^{\frac{1}{2}} - (kr)^{\frac{1}{2}}J_{n-\frac{1}{2}}(kr) \} t \\ &\quad + \{ (n+1)(kr)^{\frac{1}{2}}J_{n-\frac{1}{2}}(kr) - (kr)^{\frac{1}{2}}J_{n+\frac{1}{2}}(kr) \} / 2\sigma e^{-\sigma t} \frac{dP_n(\theta)}{d\theta}. \end{aligned} \right\} \quad (44)$$

These may be shown also to be solutions of Maxwell's equations (2) and (3). Similar solutions with $Y_{n\pm\frac{1}{2}}(kr)$ in place of $J_{n\pm\frac{1}{2}}(kr)$ also exist, but are not needed for the present problem.

To satisfy the boundary conditions these solutions multiplied by appropriate constants must be introduced as well as those previously used in the sphere and the shell. It is found that only terms of order $(n-1)$ are needed. Both in the core and in the shell the coefficient $\mathfrak{C}_{n-1}$ of the solution (44) is

$$\mathfrak{C}_{n-1} = \frac{\pi n(n+1)\omega\kappa b A}{2n+1} \frac{Y_{n-\frac{1}{2}}(ka)Y_{n-\frac{1}{2}}(kb)J_{n+\frac{1}{2}}(kb)}{J_{n-\frac{1}{2}}(ka)}.$$

The expressions for the fields may now be written down, but are somewhat cumbersome. The main point of interest is that H_ϕ contains a term proportional to

$$e^{-\sigma t} \left[J_{n-\frac{1}{2}}(kb) - \frac{2a\sigma}{b} t J_{n-\frac{1}{2}}(kb) \right] \frac{dP_{n-1}(\theta)}{d\theta},$$

where $\sigma = k^2/4\pi\kappa$ and k is given by (43). Depending on the sign of $J_{n-\frac{1}{2}}(kb)$, this may either increase or decrease with t for small t . A maximum or minimum occurs after a time

$$1 + bJ_{n-\frac{1}{2}}(kb)/2aJ_{n-\frac{1}{2}}(kb)$$

times the time of relaxation $1/\sigma$; after this time the factor $e^{-\sigma t}$ causes the term to tend monotonically to zero.

Qualitatively, the behaviour of the free modes is simple. A field that is purely toroidal decays exponentially without change of form, but a field that is initially purely poloidal produces a toroidal field as it decays. This toroidal field at first grows, but finally decays with the decay constant characteristic of the parent field. Any initial state can be represented by a suitable combination of the toroidal and poloidal solutions using, if $n > 1$ and $m = 0$, the solutions (44) in addition to those of table 1.

For $n = 1$, $m = 0$ there is no term in $n-1$, the solution (44) is not required and a simple result is obtained. The solutions of (43) are

$$k_s a = s\pi,$$

where s is an integer. This gives for the time of relaxation

$$\tau = 1/\sigma = k^2/4\pi\kappa = s^2\pi/4\kappa a^2.$$

The lowest mode has $\tau = \pi/4\kappa a^2$, a result used in II. The toroidal field in the shell is found to be given by

$$H_\phi = -\frac{1}{3}\pi^2\kappa\omega b^2 H_0 J_{\frac{3}{2}}(kb) [(s\pi - 3/s\pi) J_{\frac{3}{2}}(kr) + 3Y_{\frac{3}{2}}(kr)] (b/r)^{\frac{1}{2}} \sin 2\theta, \quad (45)$$

where H_0 is the radial field at $\theta = 0$, $r = b$. If kb is small, that is, if $a \gg \pi sb$, this gives for H_ϕ at $r = b$ and $\theta = \frac{1}{2}\pi$,

$$H_\phi = \frac{2}{3}\pi\kappa\omega b^2 H_0,$$

which is the same as the result obtained in § 6 for the stationary solution. The radial variation given by (45) is, however, different from that in (25a).

12. EXTENSIONS OF THE ANALYSIS

It would be possible to use the methods of this paper to discuss problems in which currents are induced in a rotating sphere by a varying field. Such solutions would involve no special difficulty for fields varying exponentially or simple harmonically with the time. If the assumed decay constant of the inducing field happened to coincide with the natural decay of one of the free modes, it would be necessary to use the solutions (44) as well as those of table 1. Other kinds of variation of the inducing field would involve a series of values of k or an integration over a continuous set of values. The treatment would be similar to that for a stationary sphere (Chapman & Bartels 1940).

It would also be possible, at the expense of some complication in the algebra, to take the conductivity and susceptibility of the shell as different from that of the sphere.

Similar calculations for a cylinder would be practicable, and simple problems concerning rotating ellipsoids could probably be solved.

Problems in which the angular velocity in a sphere is an arbitrary function of the radius present no formal difficulty as (6) still holds, and the variables separate in polar co-ordinates. The radial functions would no longer be Bessel functions and would have to be determined by numerical integration.

For other types of motion the right-hand side of (5) will contain θ and ϕ explicitly. It will not reduce to the simple form (6) and the variables will not separate. Such problems must be dealt with by perturbation, or other approximate methods. It is hoped that the detailed discussion of a simple problem by elementary methods that is given in the present paper will illustrate the types of solution that are to be expected, and the difficulties that may arise in the more intricate problems that must be solved in the elucidation of the origin of terrestrial magnetism.

REFERENCES

- Alfvén, H. 1947 *Mon. Not. R. Astr. Soc.* **107**, 211–219.
 Bullard, E. C. 1948 *Mon. Not. R. Astr. Soc. Geophys. Suppl.* **5**, 248–257.
 Bullard, E. C. 1949 *Proc. Roy. Soc. A*, **197**, 433–453.
 Chapman, S. & Bartels, J. 1940 *Geomagnetism*. Oxford: Clarendon Press.
 Elsasser, W. M. 1946a *Phys. Rev.* **69**, 106–116.
 Elsasser, W. M. 1946b *Phys. Rev.* **70**, 202–212.
 Elsasser, W. M. 1947 *Phys. Rev.* **72**, 821–833.

- Gans, R. 1921 *Arch. Elektrotech.* 9, 413–426.
 Lamb, H. 1881 *Proc. Lond. Math. Soc.* 13, 51–66.
 Mathematical Tables Project 1947 *Tables of spherical Bessel functions*, 1 and 2. New York: Columbia University Press.
 Thomson, J. J. 1893 *Notes on recent researches in electricity and magnetism*. Cambridge University Press.
 Tolman, R. C. 1934 *Relativity, thermodynamics and cosmology*. Oxford: Clarendon Press.
 Valentiner, S. 1927 *Handbuch der Physik*, 15.
 Watson, G. N. 1944 *A treatise on the theory of Bessel functions*. Cambridge University Press.

The flow under gravity of a swirling liquid through an orifice-plate

BY A. M. BINNIE AND J. F. DAVIDSON

Engineering Laboratory, University of Cambridge

(Communicated by Sir Richard Southwell, F.R.S.—Received 9 April 1949)

[Plate 19]

An examination has been made of the whirlpool formed when a perfect liquid possessing swirl passes under gravity through a circular sharp-edged hole in the base of a large tank. Relaxation methods were applied to determine the velocity distribution in the liquid and the shapes of the air-core and of the initial part of the emergent jet. In the numerical case considered, the swirl was large; and it was found that, except close to the orifice, the radial and vertical velocities were too small to influence the shape of the air-core within the tank. Below the orifice the liquid was emitted as a conical and rapidly thinning jet. In spite of the contraction at the edge of the hole, the discharge was not greatly different from that determined by Binnie & Hookings' (1948) approximate method for flow through a trumpet. Experiments with water were made for comparison with the calculations.

It was noticed that under certain conditions of low swirl the annular jet assumed a cylindrical vibrating form. The theory of these oscillations, which was confirmed by observation, shows that, unless the air-core occupies a large fraction of the cross-section of the jet, the wavelength is nearly the same as that in a 'solid' jet of the same outer diameter.

1. INTRODUCTION

When water is led through an orifice in the base of a tank an air-cored vortex is formed if the swirl is sufficiently great. For a perfect liquid it was shown by Binnie & Hookings (1948) that the problem of calculating the discharge under gravity can be readily solved provided that radial velocities are assumed to be negligible and that the circular orifice is trumpet-shaped, giving the jet a specified outer diameter in the plane of the orifice. In the following pages a more difficult case is considered, in which the orifice is sharp-edged and radial velocities are taken into account, so that initially the outer diameter of the emergent jet, as well as the inner, is unknown. An endeavour has been made by relaxation methods to determine the shape of the air-core and the velocity distribution in the liquid. The calculations were pursued for some distance below the orifice in order to determine the shape of the jet in that region, but no attempt was made to investigate the break-up of the jet into drops.

Experiments were carried out with water to serve partly as a guide to the calculations and partly for comparison with them. They were made on a scale sufficiently

large to ensure that the effect of surface tension upon the diameter of the air-core was negligible. A chance observation, made when the tank had almost drained down, revealed that the issuing jet could take up the annular vibrating shape shown in figure 1*a*, plate 19. In the concluding section the theory of this type of jet is examined and is compared with experiment.

2. WHIRLPOOL EXPERIMENTS

The apparatus used was simple. A brass plate, containing a sharp-edged orifice 2 in. in diameter, was inserted in the centre of the base of an open cylindrical tank, 1 ft. 6 in. high and 3 ft. in diameter. Two inlet pipes, symmetrically connected to the main supply, passed vertically up through the bottom of the tank near the ends of a diameter. They terminated in $\frac{3}{4}$ in. elbows which discharged horizontally in opposite tangential directions. The flow was smoothed and the swirl controlled by placing in the tank a number of cylindrical baffles made of perforated zinc. After leaving the orifice the water fell into a measuring tank. To determine the shape of the air-core within the cylindrical tank, a point gauge with a horizontal cross-wire was moved in a diametral plane by means of traversing gear fixed above the tank. An approximate value of the swirl at the surface was obtained by measuring the time of rotation of small floating particles. If their size were carefully chosen they would remain for several revolutions at an almost constant radius, but the method had a limited application, for at large radii the motion was unsteady and at small the velocities were too high.

Three kinds of emergent jet were noticed. When the swirl was vigorous, both surfaces of the jet showed *vena contracta* near the orifice and then expanded for a short distance before they disintegrated. The outer diameter of the *vena contracta* was measured with a theodolite horizontally sighted. With a smaller swirl the jet was spindle-shaped. The surfaces after their initial divergence coalesced on the axis lower down, both being drawn towards the axis by surface tension; thereafter the jet continued as an irregular stream. A feeble swirl produced the vibrating form shown in figure 1*a*.

For comparison with theory the air-core was measured when the depth in the tank at radius 6 in. was 13.33 in., the swirling velocity at this radius was 14.3 in./sec. and the discharge was 54.0 cu.in./sec. The radial and vertical velocities were very small at this radius, and the stagnation level was taken as $13.33 + 14.3^2/(2g) = 13.59$ in. above the orifice. The observed air-core in the region above the orifice is shown to scale by the broken line in figure 2. Under these conditions the jet was of the first type mentioned above, and the readings were not appreciably disturbed by surface waves travelling round the tank in the manner described by Binnie & Hookings.

3. THEORY OF THE WHIRLPOOL

The relaxation solution of the problem was obtained by J.F.D. The flow under consideration is symmetrical about a fixed vertical axis, and it is irrotational, being generated by gravity forces only. We take the axis of symmetry *Oz* to be directed

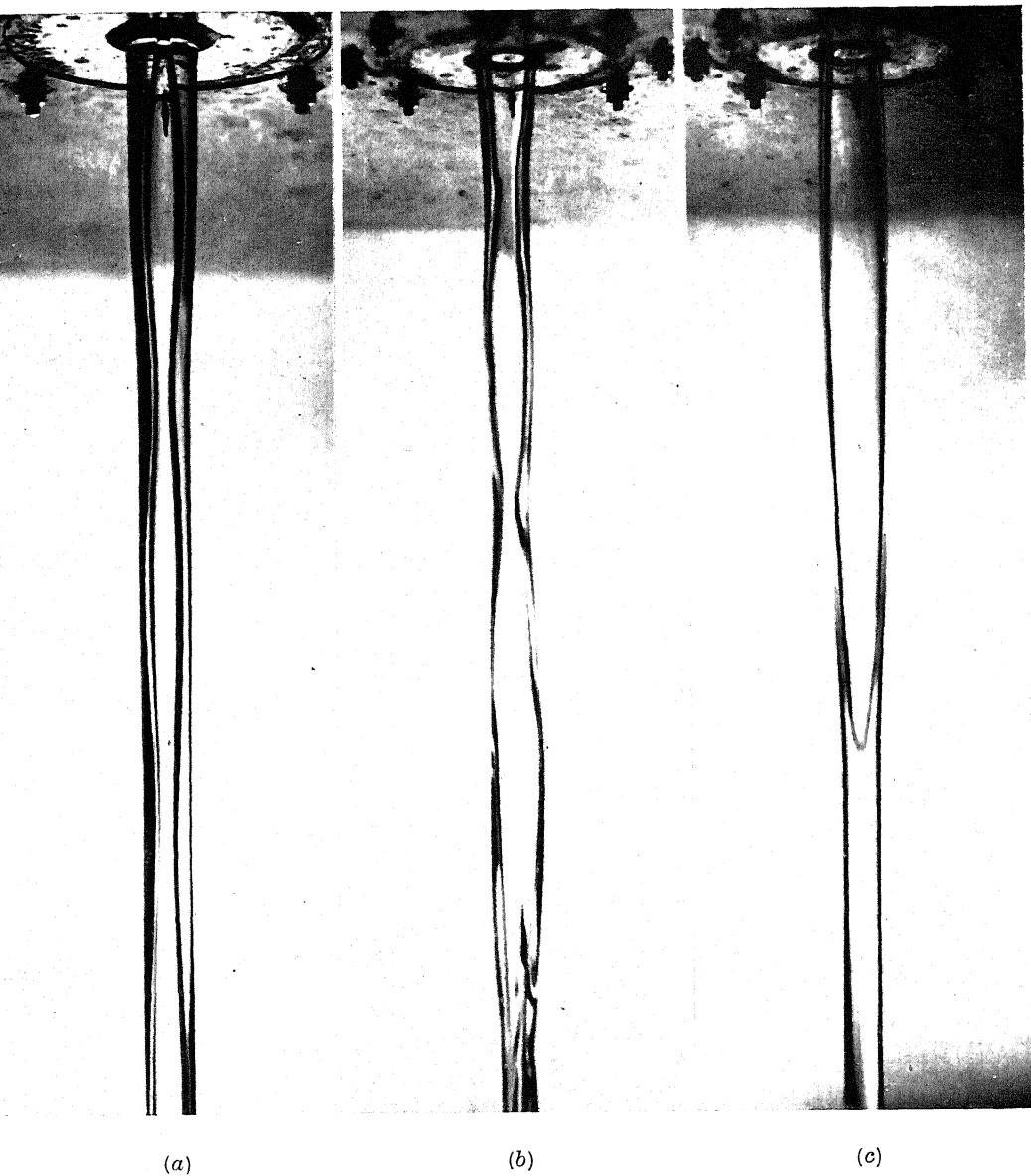


FIGURE 1. Annular jets. (a) and (b) with swirl and vibration, (c) without swirl.

downwards (figure 2) and the component velocities to be u, v, w measured positively in the direction of increasing r, θ, z . It was shown by Rayleigh (1916) that the equations of irrotational motion in the r, z plane are unaltered when a distribution of tangential velocity, given by

$$vr = \text{constant} = S \text{ say,} \quad (3.1)$$

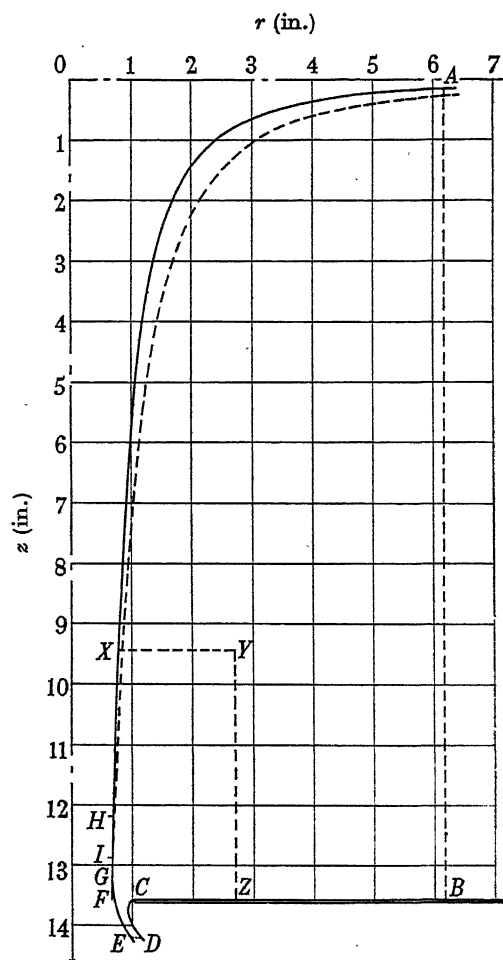


FIGURE 2. Half-sectional elevation of whirlpool. ---- observed, — calculated.

is superposed. Accordingly, we adopt (3.1); and to determine u and w we employ Stokes' stream function ψ and the well-known relations

$$u = -\frac{1}{r} \frac{\partial \psi}{\partial z}, \quad (3.2)$$

$$w = \frac{1}{r} \frac{\partial \psi}{\partial r}, \quad (3.3)$$

$$\frac{\partial^2 \psi}{\partial r^2} - \frac{1}{r} \frac{\partial \psi}{\partial r} + \frac{\partial^2 \psi}{\partial z^2} = 0. \quad (3.4)$$

The liquid being at rest at $z = 0$, we find with the aid of (3.1) that Bernoulli's equation takes the form

$$\frac{p}{\rho} + \frac{u^2 + w^2 + S^2/r^2}{2g} - z = 0, \quad (3.5)$$

in which p and ρ denote pressure and density. Hence, if ψ is given over a specified fixed boundary, the motion is unaffected by the swirl; but the pressures are decreased everywhere by $\rho S^2/(2gr^2)$, and consequently those boundaries which depend upon pressure have their positions changed.

We shall have occasion to use three types of boundary condition:

(i) The velocity perpendicular to a given boundary may be known. The difference between the values of ψ at two points P and O on the boundary can then be calculated from the relation $\psi_P - \psi_O = \int_O^P t r ds$, where t is the velocity perpendicular to the boundary of which ds is an element.

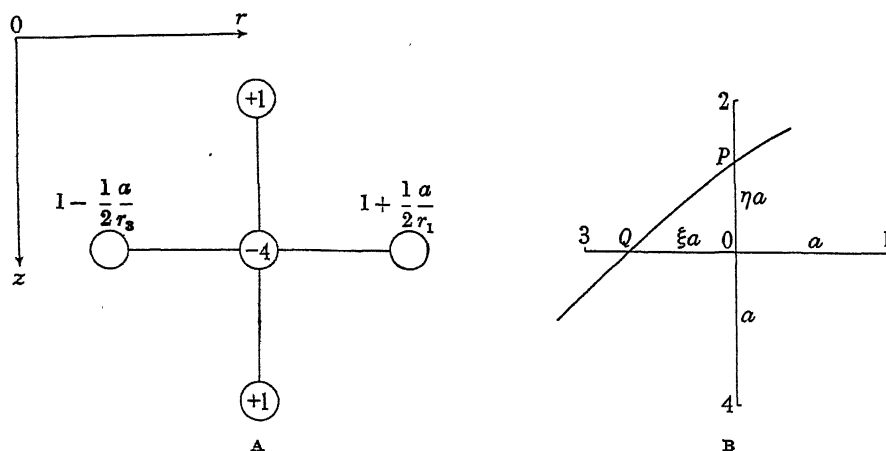


FIGURE 3. (A) Relaxation pattern, (B) intersection of mesh and curved boundary.

(ii) If the pressure p on a given boundary is specified, the magnitude (but not the direction) of the velocity $(u^2 + w^2)^{1/2}$ can be obtained from (3.5). The distribution of ψ found from (3.4) must be such as to give this velocity at every point of the boundary.

(iii) Should the boundary be a free surface, its position is not initially given. Over it, not only must (3.4) be satisfied, but also the two further conditions that the values of ψ and p are known constants.

An account was given by Southwell (1946) of the relaxational treatment of the governing equation (3.4), and details of the application of the method to free-surface problems not involving swirl were published by Southwell & Vaisey (1946). A finite difference approximation, given by their equation (22), replaces the governing equation and leads to definitions of 'residual' and 'relaxation pattern'. The pattern appropriate to a square mesh of side a is displayed in figure 3A. Where a mesh line crosses a curved boundary, as shown in figure 3B, the method of calculating the residuals at neighbouring points is altered. After ψ_1 , ψ_P , ψ_Q and ψ_4 have been ex-

panded by means of Taylor's theorem in terms of ψ and its derivatives at the point 0, it may easily be proved that the residual at 0 is

$$\begin{aligned} \psi_1 \left[\frac{2 \left(1 - \frac{\xi a}{2r_0} \right)}{1 + \xi} \right] + \psi_4 \left[\frac{2}{1 + \eta} \right] + \psi_Q \left[\frac{2 \left(1 + \frac{a}{2r_0} \right)}{\xi(1 + \xi)} \right] \\ + \psi_P \left[\frac{2}{\eta(1 + \eta)} \right] - \psi_0 \left[\frac{2}{\xi} + \frac{2}{\eta} + \frac{(1 - \xi) \frac{a}{r_0}}{\xi} \right]. \quad (3.6) \end{aligned}$$

Thus new patterns must be formed for points 0, 1 and 4. It should be noted that the finite difference equations, and therefore the residuals, at points 1 and 4 are unaltered.

When ψ is known along the boundary as in condition (i), relaxation patterns can be formulated for all points in the manner set out above, and the solution is then straightforward. For boundaries with condition (ii) the approach is more tentative. The distribution of velocity perpendicular to the boundary must first be guessed and the resultant problem solved as in (i). From this solution the total velocity and hence the pressure at each point on the boundary can be found. The hypothetical distribution of perpendicular velocity must now be altered to bring the pressures closer to their specified values, and the process is repeated until the two are the same. For condition (iii) the procedure is similar except that initially the shape of the boundary must be guessed. Since ψ is given along it, the velocity and pressure distribution may be obtained, and the shape is then adjusted until the pressure is correct.

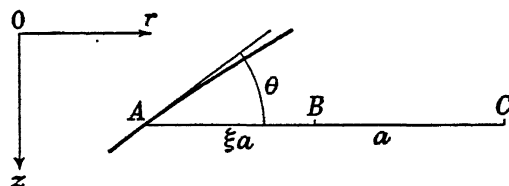


FIGURE 4. Determination of q on boundary.

In the last two cases it is necessary to find the velocity $q = (u^2 + w^2)^{1/2}$ along the boundary at a point A (figure 4) when ψ is known at A and has been calculated at B and C by the relaxation process. Taylor's theorem is used to give expansions for ψ_B and ψ_C in terms of ψ , $\partial\psi/\partial r$ and $\partial^2\psi/\partial r^2$ at A . The considerations set out in Southwell & Vaisey's § 15 then yield the result

$$q_A = \frac{1}{ar_A \sin \theta} \left[\frac{1 + \xi}{\xi} \psi_B - \frac{\xi}{1 + \xi} \psi_C - \left\{ \frac{1 + 2\xi}{\xi(1 + \xi)} \right\} \psi_A \right]. \quad (3.7)$$

The problem in hand can now be solved by means of the methods described above. The boundary conditions are as follows:

(a) Along $AXGE$ (figure 2) we have the double condition $\psi = \text{constant}$, arbitrarily chosen as zero,

$$\frac{u^2 + w^2 + S^2/r^2}{2g} - z = 0, \quad (3.8)$$

the latter equation being obtained from (3.5).

(b) Along AB , where $r = 6.175$ in., u was taken as constant—an assumption which at this large radius seems reasonable.

(c) Along BC , $\psi = \text{constant} = k$, where k is one of the results of the computations and must therefore be guessed initially.

(d) Along CD , $\psi = k$ and (3.8) holds good.

(e) Along DE where the stream is very thin the pressure was taken as constant. This assumption was made for positions of DE at successively increasing distances from C until the change in the free surfaces caused by a move of DE was negligible. This procedure is justified by the view that if the liquid were perfect the sheet would not disintegrate.

To facilitate experimental checks the calculations were carried out in inch-second units. As far as possible they were based on the experiments, therefore the orifice diameter was taken as 2 in. and the stagnation level as 13.59 in. It was more difficult to decide on the value of the swirl constant S . Its observed value at $r = 6$ in. was $85.8 \text{ in.}^2/\text{sec.}$, and it was found that, along the upper part of the measured inner surface AG (figure 2), $u^2 + w^2$ was negligible, therefore in accordance with (3.8) this curve was almost identical with that defined by $z = 85.8^2/(2gr^2)$. The latter, when extended down to the orifice, showed an inner diameter appreciably greater than that observed, because, although v remained the predominant component of velocity, the effective value of S was diminished by viscous forces. With these considerations in mind a smaller value $S = 66.9 \text{ in.}^2/\text{sec.}$ was used in the calculations, the curve $z = 66.9^2/(2gr^2)$ being judged the best fit with the observations at the orifice.

In order that a start could be made, the boundary AGF (figure 2) was taken as $z = 66.9^2/(2gr^2)$, and to estimate the discharge the pressure and radial velocity across FC were taken as zero. By integration the discharge Q is then given by

$$Q = \frac{\pi S^2}{(2gH)^{\frac{1}{2}}} \left[\frac{r_2}{r_1} \left(\left(\frac{r_2}{r_1} \right)^2 - 1 \right)^{\frac{1}{2}} - \cosh^{-1} \frac{r_2}{r_1} \right], \quad (3.9)$$

where H is the total head, and r_1 and r_2 are the radii at F and C . This approximation is clearly an over-estimate. The use of boundary condition (b) made it possible to calculate the distribution of ψ over AB , and thence the stream-lines in the tank were obtained by the relaxation process with a square mesh of side $a = 0.690$ in. Some of them were far from being perpendicular to AB , showing that the solution cannot claim to be valid for a very large tank, but the conditions at AB , which were doubtful in the experiments, do not greatly influence the flow close to the orifice. Next, the velocity q along AF was calculated at each point by means of (3.7). It should be noted that because the discharge was over-estimated these values were too great.

Now if, in order to satisfy the pressure condition (3.8), the radius of the original boundary, over which $v^2 = 2gz$, is increased by a small amount dr , the resultant change in q will be very small. It follows from (3.8) that the corresponding change dv in the tangential velocity is given by

$$q^2 + (v + dv)^2 = 2gz = v^2, \quad (3.10)$$

and hence

$$dv = -\frac{q^2}{2v}. \quad (3.11)$$

But from (3.1)
$$dv = -\frac{S}{r^2}dr, \quad (3.12)$$

therefore
$$\frac{dr}{r} = \frac{q^2}{2v^2}. \quad (3.13)$$

In this way the change in r required to satisfy the pressure condition may be calculated. For the three points F , I , H in figure 2 immediately above the orifice, these changes were found to be 1.77, 0.20 and 0.054 %. As the largest is so small, it appears that in the powerful vortex under consideration $z = 66.9^2/(2gr^2)$ is a very close approximation to the shape of the air-core above the orifice-plate.

The second mesh with side $a = 0.345$ in. revealed no change in ψ at points farther from the orifice than XYZ (figure 2), therefore the third and fourth, for which a was 0.1725 and 0.08625 in., were confined to the region below XYZ . For these the values of ψ required along XYZ were obtained from the second mesh by graphical interpolation. In view of the doubtful starting assumption along AB (figure 2), the precise method of interpolation employed by Southwell & Vaisey was deemed unnecessary. In the fourth mesh a refinement was made to the conditions in the neighbourhood of the orifice. The core was kept on the curve $z = 66.9^2/(2gr^2)$ which has been shown to be accurate to 2 % in this region. The outer surface was guessed as the curve CR (figure 5), the point R being at the *vena contracta*, and across a horizontal section through R the pressure and radial velocity were assumed to be zero, so that the discharge could be estimated from (3.9). The values of ψ , q and dr/r were then determined as before. For points on the core the values of $2dr/a = dr/0.0431$ are shown in the column on the left of figure 5, from which it can be seen that the change in r is small in the region LO but larger close to S . This figure gives the final values of ψ for the upper part of the field, but, the available space being limited, only a quarter of the values actually obtained can be inserted. The mesh size was too big to yield satisfactory values of q along CR .

To obtain these values, the fifth mesh ($a = 0.043125$ in.) was used inside the region OPQ (figure 5), and it was found that the values of ψ outside GJK were unaffected by changes near the orifice. Most of the work of finding the shapes of the outer free surface and of the lower part of the core was done on this mesh, as it gave reasonably accurate values of q with the curvatures used, and the mesh was not so small as to make the calculations excessively laborious. For a start, the values of ψ had been determined along $GJKC$, and GF (figure 6) is the curve $z = 66.9^2/(2gr^2)$ which was known to be coincident with the core above G . The core was guessed as GeU and the outer free surface as CdT . The pressure across TU was taken to be constant, and the angle between the stream-lines and TU to be a linear function of r . The discharge was then calculated by graphical integration across TU , and again the distribution of ψ was obtained by the relaxation method and thence the values of q on the two free surfaces; these values had to be made consistent with (3.8). The process was assisted by plotting (3.8) as graphs of q against r for various values of z , which changed only slightly in the region under consideration. Thus it was necessary to discover shapes for CdT and GeU such that, at every point, q as determined by the relaxation

the errors in the calculated values of q being shown by ringed figures. For the reason pointed out above, it is difficult to assess the accuracy of the solution in terms of errors in radius, but comparison with previous attempts suggests that the final result

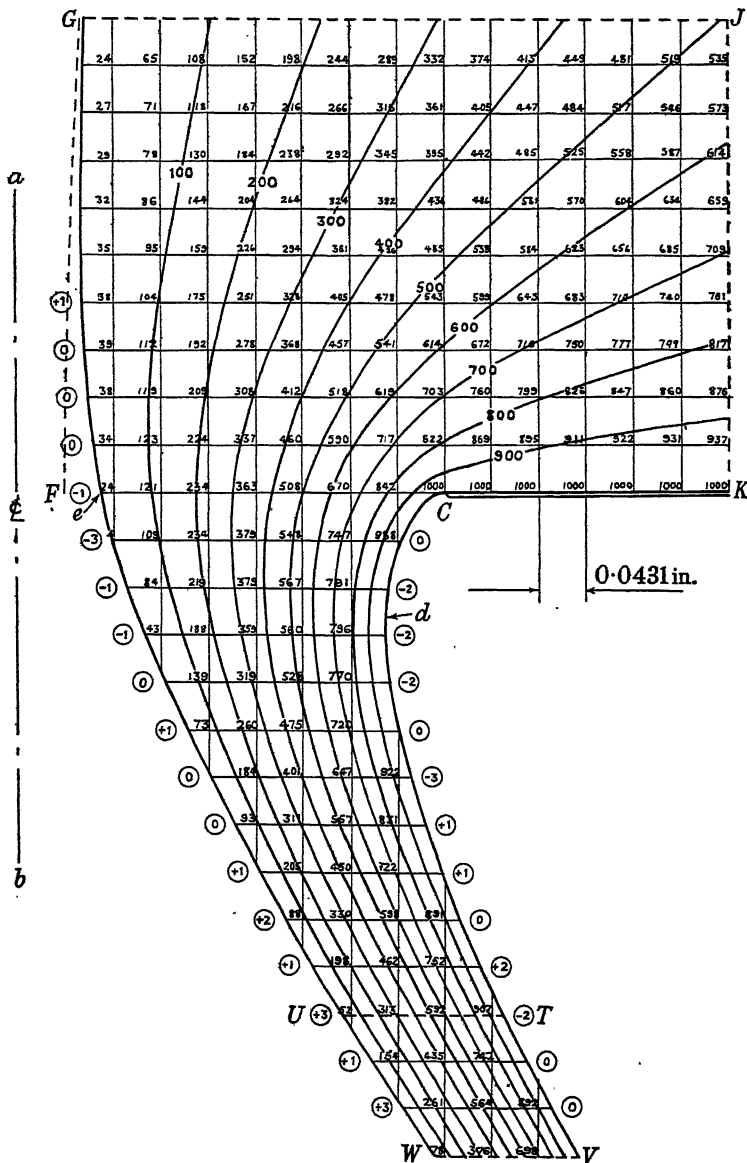


FIGURE 6. Values of ψ in lower part of field. The axis ab is displaced to the right by a distance (0.612 in.) equal to the length ab . The errors in the values of q are shown by ringed figures.

is correct to 0.005 in. This limit is, of course, far less than the errors in the experiments, where the measurements were disturbed by surface-tension ripples.

The calculated *vena contracta* of the outer free surface had a radius 0.946 in. as against the measured value of 0.909 in., and the calculated and experimental discharges were 63.8 and 54.0 cu.in./sec. These differences, due to boundary-layer

effects, are in accordance with normal experience, but it must be remembered that the value ($66.9 \text{ in.}^2/\text{sec.}$) of S employed in the calculations was less than the observed entry value ($85.8 \text{ in.}^2/\text{sec.}$). Had the latter been used, the flow through the orifice would have been confined to a smaller annulus and the discharge decreased to below the experimental value. Thus, as was pointed out by Binnie & Hookings, the effect of viscosity is to reduce the size of the core and consequently to increase the discharge.

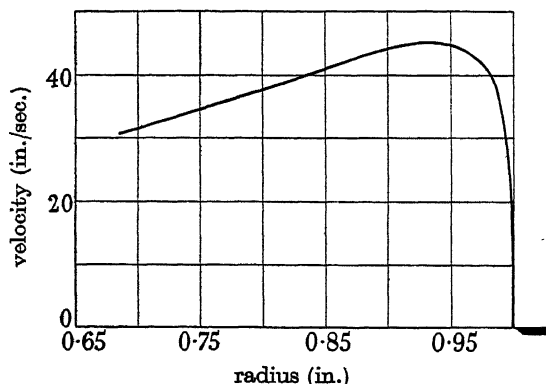


FIGURE 7. Downward velocity in plane of orifice.

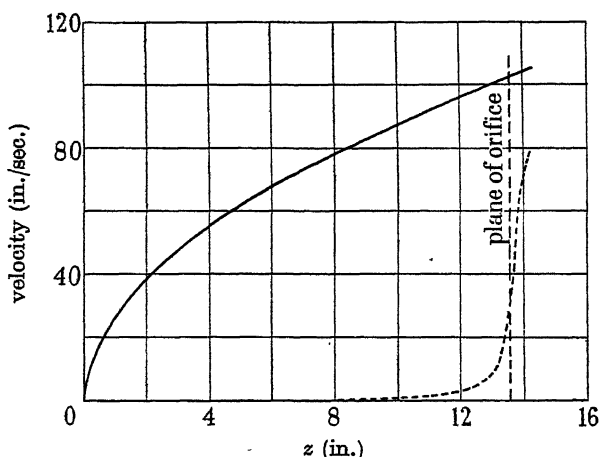


FIGURE 8. Velocities on surface of core. — total velocity, ---- $q = (u^2 + w^2)^{1/2}$.

Figure 7 shows how the calculated downward velocity w varies over the plane of the orifice. This velocity is zero at the edge; it rises rapidly to a maximum at radius 0.93 in. and then decreases slowly to the surface of the core; thus it may be regarded as nearly constant. The approximate method used by Binnie & Hookings, it is interesting to note, predicts the discharge with surprising accuracy, although radial velocities are far from small and the conditions at the outer radius are widely different. From the assumption that radial velocities are negligible they deduced that w is constant over the plane of the orifice. On equating to zero the derivative of the discharge with respect to the unknown inner radius, they obtained expressions for the values of this radius and of the discharge. When applied to the present case

their formulae yield the results 0.763 in. and 68.5 cu.in./sec., which should be compared with 0.686 in. and 63.8 cu.in./sec. obtained by the relaxation process. Finally, the variation of velocity over the core is given in figure 8, in which the full line shows the total velocity and the broken line the values of $q = (u^2 + w^2)^{\frac{1}{2}}$. The latter indicates that the downward velocity is very small until the orifice is approached, showing why a light particle placed on the surface descends at first with extreme slowness.

4. THE ANNULAR SWIRLING JET

It was noticed that, if the tank were allowed to drain down when the swirl was small, the issuing jet could assume the annular vibrating form shown in figure 1*a*, plate 19. This type of jet is analogous to the vibrating 'solid' jet investigated by Rayleigh (1878, 1879), whose work was extended by Pedersen (1907) and by Bohr (1909) to provide an accurate means of determining the surface tension of the newly formed and therefore uncontaminated surface. In the non-rotating annular jet, which was examined by Binnie & Squire (1941), surface tension acting on the two boundaries causes them to coalesce some distance from the orifice; thereafter the jet has the 'solid' form as shown in figure 1*c*, plate 19. But if the jet is spinning, so that the surface-tension forces are balanced by the centrifugal, vibration about the equilibrium position is possible. In the following paragraphs the theory of this jet is developed and is compared with experiment.

The first step is to examine the equilibrium of the jet under the action of the centrifugal and surface-tension forces. The complications due to the effect of gravity will not be taken into account; and as the velocity in the direction of the jet axis is assumed to be uniform, this also may be disregarded in the analysis. Accordingly it is sufficient to take a cross-section of an infinite cylinder, the outer and inner radii being denoted by a and b . Since the motion is irrotational, the tangential velocity v at radius r is given by $v = S/r$, S being constant throughout the jet. We consider an element of inner radius r , radial thickness dr , length unity, and subtending a small angle $d\theta$ at the centre. This is in radial equilibrium provided that

$$\left(p + \frac{dp}{dr} dr\right)(r + dr) d\theta - pr d\theta - p dr d\theta = \rho \frac{S^2}{r^2} dr d\theta, \quad (4.1)$$

where ρ is the density of the liquid and p is the excess of pressure above that of the surrounding atmosphere. On integration (4.1) reduces to

$$p = \text{constant} - \frac{\rho S^2}{2r^2}. \quad (4.2)$$

The surface conditions are

$$p = T_1/a \quad \text{at} \quad r = a, \quad p = -T_1/b \quad \text{at} \quad r = b, \quad (4.3)$$

where T_1 is the surface tension. The insertion of (4.3) into (4.2) then yields

$$\frac{a-b}{ab} = \frac{2T_1}{\rho S^2} \quad (4.4)$$

as the necessary relation between a and b when S and the properties of the liquid are specified.

We proceed now to determine the nature of the equilibrium of the jet. For this purpose it is convenient to extend the compact method used by Lamb (1932) in the analysis of the 'solid' jet rather than the longer energy method employed by Rayleigh in his original investigation. The motion being generated from rest, it may be assumed that a velocity potential ϕ exists; and, since the liquid is taken to be inviscid, in cylindrical co-ordinates (r, θ, z) ϕ must satisfy

$$\frac{\partial^2 \phi}{\partial r^2} + \frac{1}{r} \frac{\partial \phi}{\partial r} + \frac{1}{r} \frac{\partial^2 \phi}{\partial \theta^2} + \frac{\partial^2 \phi}{\partial z^2} = 0. \quad (4.5)$$

The appropriate form of solution giving symmetry about the axis is

$$\phi = \phi_1 \cos kz \cos(\sigma t + \epsilon) - S\theta, \quad (4.6)$$

in which ϕ_1 is a function of r only. It will be seen that (4.6) gives the correct distribution of tangential velocity, for

$$v = -\frac{1}{r} \frac{\partial \phi}{\partial \theta} = \frac{S}{r}.$$

On substituting (4.6) in (4.5) we obtain

$$\frac{\partial^2 \phi_1}{\partial r^2} + \frac{1}{r} \frac{\partial \phi_1}{\partial r} - k^2 \phi_1 = 0, \quad (4.7)$$

of which the solution is $\phi_1 = AI_0(kr) + BK_0(kr), \quad (4.8)$

where A and B are constants. It is now supposed that the outer radius is

$$r = a + \xi, \quad (4.9)$$

in which ξ represents a vibration of infinitely small amplitude. Since

$$\frac{\partial \xi}{\partial t} = -\frac{\partial \phi}{\partial r} \quad \text{at} \quad r = a,$$

we find from (4.6) and (4.8) that

$$\xi = -\frac{k}{\sigma} \{AI_1(ka) - BK_1(ka)\} \cos kz \sin(\sigma t + \epsilon). \quad (4.10)$$

It was shown by Lamb that at the outer radius the sum of the principal curvatures is

$$\frac{1}{R_1} + \frac{1}{R_2} = \frac{1}{a} - \frac{1}{a^2} \left(\xi + \frac{\partial^2 \xi}{\partial \theta^2} \right) - \frac{\partial^2 \xi}{\partial z^2}, \quad (4.11)$$

hence the pressure there is found from (4.10) to be

$$\begin{aligned} p &= T_1 \left(\frac{1}{R_1} + \frac{1}{R_2} \right) \\ &= T_1 \left[\frac{1}{a} - \frac{k}{\sigma a^2} (k^2 a^2 - 1) \{AI_1(ka) - BK_1(ka)\} \cos kz \sin(\sigma t + \epsilon) \right]. \end{aligned} \quad (4.12)$$

Another expression for the pressure can be derived from Bernoulli's equation

$$\frac{p}{\rho} = \frac{\partial \phi}{\partial t} - \frac{1}{2} q^2 \quad (4.13)$$

(where q is the resultant velocity), which together with (4.6) and (4.8) yields

$$p = -\rho \left[\sigma \{A I_0(ka) + B K_0(ka)\} \cos kz \sin(\sigma t + \epsilon) + \frac{1}{2} \frac{S^2}{a^2} \right]. \quad (4.14)$$

We are concerned only with the terms in (4.12) and (4.14) which are variable with time; the remainder reduce to a constant in virtue of (4.2) and (4.3). Hence

$$\sigma^2 = T_1 \frac{ka}{\rho a^3} (k^2 a^2 - 1) \left\{ \frac{I_1(ka) - x K_1(ka)}{I_0(ka) + x K_0(ka)} \right\}, \quad (4.15)$$

in which, for brevity, x is written for B/A .

The motion of the inner surface may be dealt with in the same manner. Denoting the inner radius by

$$r = b + \eta, \quad (4.16)$$

we obtain
$$\eta = -\frac{k}{\sigma} \{A I_1(kb) - B K_1(kb)\} \cos kz \sin(\sigma t + \epsilon). \quad (4.17)$$

But now the surface tension produces a negative pressure at $r = b$. Hence the equation corresponding to (4.15) is

$$\sigma^2 = T_1 \frac{kb}{\rho b^3} (1 - k^2 b^2) \left\{ \frac{I_1(kb) - x K_1(kb)}{I_0(kb) + x K_0(kb)} \right\}, \quad (4.18)$$

and the right-hand sides of (4.15) and (4.18) taken together provide a means of determining x . The jet will be unstable if σ^2 as found from (4.15) is negative, and (in accordance with the argument developed by Rayleigh) the preferred wave-length of the vibration will be given by that value of ka which makes

$$Z = ka(1 - k^2 a^2) \left\{ \frac{I_1(ka) - x K_1(ka)}{I_0(ka) + x K_0(ka)} \right\} \quad (4.19)$$

a maximum. An identical result is, of course, obtained if the same reasoning is applied to (4.18).

To make the above results intelligible, detailed calculations have been carried out for two cases:

(i) For $b/a = 0.15$, corresponding very approximately to figure 1*a* (in which, owing to the lens effect of the outer surface, the ratio of the true to the apparent inner diameter is about three-quarters). The cross-sectional area enclosed by the inner circumference is only about one-fiftieth of that enclosed by the outer, hence it was thought desirable to work out a second example, which corresponds more closely to figure 1*b*.

(ii) For $b/a = 0.6$. Here the area ratio mentioned above is about one-third.

The relation between ka and the wave-length λ of the undulations is

$$\lambda = \pi \frac{2a}{ka}, \quad (4.20)$$

and observations of the jet invariably showed wave-lengths exceeding π times the outer diameter. Accordingly, the numerical explorations were confined to values of ka less than unity.

The quadratic equation for x which emerges from the combination of (4.15) and (4.18) was found always to have two real roots, one large and negative and the other small and positive. These two cases will be dealt with in turn.

(1) x negative. Since the Bessel functions are essentially positive, it is seen at once from (4.10) and (4.17) that, as long as x is negative, ξ and η have the same sign. Thus we are investigating the type of jet in which the outer and inner surfaces vibrate in phase. The values of σ^2 were all negative but, as indicated in table 1, no maximum value of $Z^\dagger$ appeared. We conclude that this type is theoretically unstable and does not exhibit a preferred wave-length.

TABLE 1. THE CRITERION OF INSTABILITY $Z^\dagger$ FOR NEGATIVE
VALUES OF x

ka	$b/a = 0.15$	$b/a = 0.6$
0.05	—	3.318
0.1	—	3.312
0.4	12.60	3.197
0.6	12.67	3.038
0.7	12.72	—
0.8	12.78	2.798
0.9	12.84	2.642

(2) x positive. The signs of ξ and η cannot in this case be determined by inspection, but the calculations showed that they were always opposite, therefore the two surfaces vibrate in anti-phase. Again σ^2 was found to be negative, but in this case $Z^\dagger$ shows a maximum value. The details are given in table 2, to which Rayleigh's results for the 'solid' jet have been added for purposes of comparison. From the figures given in table 2 the values of ka corresponding to the maxima of $Z^\dagger$ were

TABLE 2. THE CRITERION OF INSTABILITY $Z^\dagger$ FOR POSITIVE
VALUES OF x

ka	$b/a = 0$ (Rayleigh)	$b/a = 0.15$	$b/a = 0.6$
0.4	0.2567	0.2536	0.1895
0.5	0.3015	0.2981	0.2250
0.6	0.3321	0.3285	0.2511
0.7	0.3433	0.3398	0.2638
0.8	0.3269	0.3238	0.2563
0.9	0.2647	—	0.2128

determined by trial using Everett's interpolation formula as far as second differences. The conclusions, together with the corresponding wave-lengths, are set out in table 3, which includes the result given by Rayleigh. The theory therefore indicates that a jet of this type will develop undulations of wave-length about $4\frac{1}{2}$ outer diameters as long as the central core is small. The wave-length will fall only slightly when the ratio b/a is considerably increased.

TABLE 3. THE PREFERRED WAVE-LENGTHS OF VIBRATION

b/a	ka for max. $Z^{\frac{1}{2}}$	$\lambda/(2a)$
0 (Rayleigh)	0.696	4.51
0.15	0.697	4.51
0.6	0.732	4.29

The conditions favourable to the production of a vibrating jet were found difficult to maintain, and the orifice diameter was reduced to $1\frac{3}{32}$ in. in order to facilitate the experiments by increasing the surface-tension effects. At heads as high as 7 in., jets of the required type could then be formed which, when they impinged on the bottom of an empty measuring tank, emitted a loud drumming sound, thus confirming their varicose nature. But their surfaces were slightly rippled, and jets more suitable for photography were obtained at very low heads. That shown in figure 1*a* was produced at a head of about 1 in.; it corresponds very roughly to the theoretical case $b/a = 0.15$ analyzed above. The wave-length is seen to be rather more than the theoretical $4\frac{1}{2}$ outer diameters, but this is to be expected, first, because the amplitude of the vibration is not infinitely small and, secondly, because the jet velocity is not constant but is increasing.

To enable jets with thinner walls to be examined and to make the initial anti-phase conditions more definite, a thin disk of diameter $\frac{5}{8}$ in. mounted transversely on a long hollow rod was lowered centrally into the plane of the orifice. The jets thus produced were of the same form as before, although not so glassy or stable, owing no doubt to the boundary layer developed on the rod and on the upper surface of the disk. An example taken at a head of about 3 in. is shown in figure 1*b*; its dimensions agree approximately with the second theoretical case which was considered, and again the theory is roughly corroborated. For purposes of comparison a non-rotating jet produced at 13 in. head with the same apparatus is shown in figure 1*c*; for this the disk was set lower to give the jet a small initial divergent motion.

Finally, an attempt was made to produce a spinning jet in which the walls were vibrating in phase, and for this purpose a hollow rod of outer diameter $\frac{5}{8}$ in. was placed centrally in the tank, its base being nearly in the plane of the orifice. It was seen that this arrangement gave both surfaces an initial inward direction, but the inner surface continued to move inwards after the outer had begun to diverge, and thus anti-phase motion set in of its own accord in agreement with the theoretical prediction.

REFERENCES

- Binnie, A. M. & Hookings, G. A. 1948 *Proc. Roy. Soc. A*, **194**, 398.
 Binnie, A. M. & Squire, H. B. 1941 *Engineer, Lond.*, **171**, 236.
 Bohr, N. 1909 *Phil. Trans. A*, **209**, 281.
 Lamb, Sir H. 1932 *Hydrodynamics*, p. 472. Cambridge University Press.
 Pedersen, P. O. 1907 *Phil. Trans. A*, **207**, 341.
 Rayleigh, Lord 1878 *Proc. Lond. Math. Soc.* **10**, 4 (*Papers*, **1**, 361).
 Rayleigh, Lord 1879 *Proc. Roy. Soc.* **29**, 71 (*Papers*, **2**, 377).
 Rayleigh, Lord 1916 *Proc. Roy. Soc. A*, **93**, 148 (*Papers*, **6**, 447).
 Southwell, R. V. 1946 *Relaxation methods in theoretical physics*. Oxford: Clarendon Press.
 Southwell, R. V. & Vaisey, G. 1946 *Phil. Trans. A*, **240**, 117.

The transformation of the nitrogen isotope of mass 15 into carbon and helium by bombardment with protons

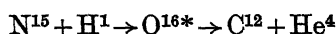
BY W. COCHRANE AND A. G. HESTER, *Imperial College of Science and Technology*

(Communicated by Sir George Thomson, F.R.S.—Received 9 May 1949)

Measurements of the 'cross-section' of the nuclear reaction $N^{15}(p, \alpha)C^{12}$ have been made for protons with different values of the energy in the interval between 200 and 450 keV. There are indications of a broad maximum, in the curve showing the variation of 'cross-section' with energy, in the neighbourhood of 400 keV, with a width of 150 keV. The partial proton width of the maximum is 250 eV. Extrapolation of the one-level dispersion formula, together with the theory of thermonuclear reactions, gives a lifetime of 3×10^4 years for N^{15} in the sun. A photographic method of determining the angular distribution of the emitted α -particles is described. The angular distribution near 400 keV is found to be approximately isotropic, the deviations from isotropy being, at most, 10 %.

1. INTRODUCTION

The α -particles from the nuclear reaction



were first detected by Burcham & Smith (1939). They obtained their result using ordinary nitrogen, which contains only 0.38 % of the heavy isotope, and their results suggested that the yield of the reaction is relatively high. This was confirmed by Holloway & Bethe (1940), who used nitrogen enriched to 70 % in N^{15} ; they measured the effective cross-section, σ , at a proton energy of 360 keV and found σ to be 1.3×10^{-26} cm.².

Apart from its intrinsic interest, the reaction $N^{15}(p, \alpha)C^{12}$ is important in forming one of the links in the nuclear carbon cycle, suggested by Bethe (1939) as an energy source in matter at temperatures of the order of 10^7 °K, such as exist in the stars. Although in matter at such a temperature, protons will possess an average energy of only a few thousand electron volts, the determination of the cross-section over a range of somewhat higher energies may provide data for extrapolation to the stellar conditions. In the present work, this variation of the cross-section of the reaction has been measured between 200 and 450 keV; and, in addition, the angular distribution of the emitted α -particles has been investigated.

2. EXPERIMENTAL ARRANGEMENTS FOR MEASUREMENT OF CROSS-SECTION

The target consisted of ammonia gas in which the nitrogen was enriched to 31 % in the isotope N^{15} . The enriched material was purchased in the form of ammonium nitrate with the additional isotope content in the ammonium radical. This salt was all converted to ammonium chloride using Kjeldahl's method. Small amounts of the ammonium chloride were then converted to ammonia, when required, by mixing with excess quicklime and heating.

The protons bombarded the ammonia gas in a target chamber which was a cavity, *A* (figure 1), of about 1 cm.³ cut out of a brass block. The protons entered the chamber through a mica window *B*, of diameter 6.35 mm., the direction of the proton beam being perpendicular to the window. The α -particles arising from disintegrations in *A* were observed in a direction at right angles to the proton beam; they issued from the chamber *A* through a mica window *C* of diameter 6.35 mm., and then travelled

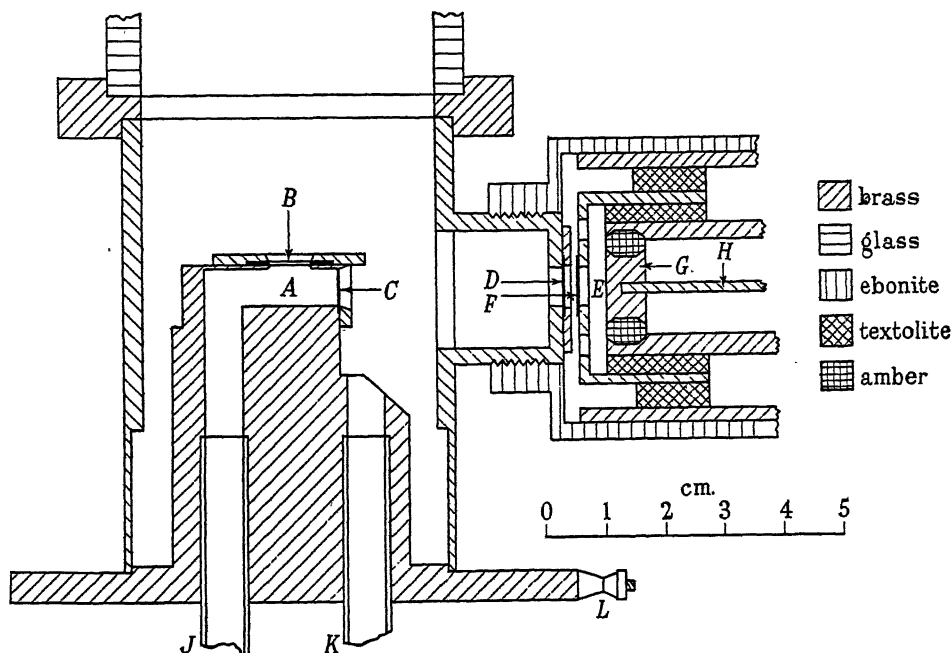


FIGURE 1. Gas target and ionization chamber. *A*, target volume; *B*, mica window with wire gauze; *C*, mica window; *D*, mica window; *E*, ionization chamber; *F*, aluminium foil; *G*, collecting electrode; *H*, brass rod; *J*, tube for introduction of ammonia; *K*, tube for equalizing pressure on both sides of micas, when required; and *L*, terminal for micro-ammeter reading proton current.

through the vacuum to the wall of the apparatus, finally emerging into the laboratory atmosphere through a second window *D* of the same diameter. The proton window *B* consisted of a thin piece of mica covered over with a fine-meshed wire gauze. When gas was admitted to the target chamber, the mica was pressed out against the wire gauze. Much of the heat generated by the protons in passing through the mica was carried off by the wires of the gauze in close contact with the mica surface. The proton windows used had a thickness equivalent to 3 mm. (or thereabouts) of atmospheric air; the first α -particle window *C* had a thickness of 3.6 mm. of air, and the second *D* 5.6 mm. of air. The pressure of ammonia in the target chamber *A* was of the order of 2 cm. of mercury. Proton current was measured from the terminal *L*. It was verified, by bringing a magnet near, that the readings were not being falsified by the escape of secondary electrons.

The α -particles were detected by a shallow ionization chamber *E*; ions formed during passage of the α -particles through the air space *E* between the aluminium

foil F and the collecting electrode G were swept on to G by an electric field of 800 V/cm. and led by the rod H to the grid of a cathode-follower valve. This was followed by a linear amplifier and recorder, as described by Wynn-Williams (1936). The recorder had a resolving time of approximately 0.01 sec. Electrical screening of chamber and amplifier was meticulously carried out, and precautions were taken against chamber and valve microphonics, so that no extraneous or 'natural' counts occurred. The counting apparatus (removed from the target position) was also employed, in conjunction with a weak polonium source, to determine the equivalent air thickness of the various mica windows.

The accelerating potential applied to the protons was generated by a van der Graaff machine designed by Smee (1944). The arrangement of ion source and accelerating tube was that described by Craggs (1942). After magnetic separation of the ion beam and introduction of collimating slits, proton currents of a few microamperes were available up to 550 kV. Fluctuations in the voltage were at most of the order of 10 kV.

3. RESULTS FOR THE CROSS-SECTION

The emitted α -particles were counted over periods ranging from 1 to 4 min. for different values of the proton accelerating voltages. After several preliminary runs which indicated the trend of the results, the data shown in table 1 were obtained. In calculating the cross-section σ from the third column of table 1 an isotropic angular distribution of the α -particles was assumed. Other figures necessary for the calculation were as follows: fraction of complete sphere subtended by ionization chamber at target = $\frac{1}{850}$; length of proton path effective in producing detectable disintegrations = 0.64 cm.; factor to allow for part of measured proton current being stopped by gauze = 0.25; air equivalent thickness of proton window = 2.1 mm.; and pressure of 31 % enriched ammonia = 1.5 cm. of mercury. The reduction in proton energy when passing through the window was calculated from the data given by Livingston & Bethe (1937). The results for σ are presented graphically in figure 2, which also shows the single observation of Holloway & Bethe (1940).

TABLE 1. EXPERIMENTAL RESULTS FOR σ

van der Graaff voltage (kV)	proton energy after window (keV)	counts per min. per μ A	$\sigma \times 10^{28}$ (cm. ²)	statistical error in counting (%)
359	205	0.4	0.3	large
375	228	1.9	1.6	30
390	252	5.8	4.9	15
422	290	28	23.5	15
437	312	66	55	12
454	335	108	90	6
469	353	116	97	7
485	375	127	105	4.5
500	393	172	144	7
516	412	126	105	6
532	430	108	90	7
550	452	105	88	6

Considering the uncertainty in reduction of energy by the mica window used here and the aluminium window used by Holloway & Bethe, the agreement between the two results is good. The standard deviations arising from the statistical nature of the counts are indicated in the last column of table 1 and on figure 2.

A rough determination of the range of the α -particles was made. A Q -value (or nuclear energy release) of 5.2 MeV was obtained. This compares with a theoretical Q -value of 5.0 MeV based on the latest mass-spectroscopic data, but the difference is not considered significant. An accurate determination of range would have required a more elaborate technique. The two values are sufficiently close, however, to identify the α -particles as corresponding to the full Q -value of the reaction.

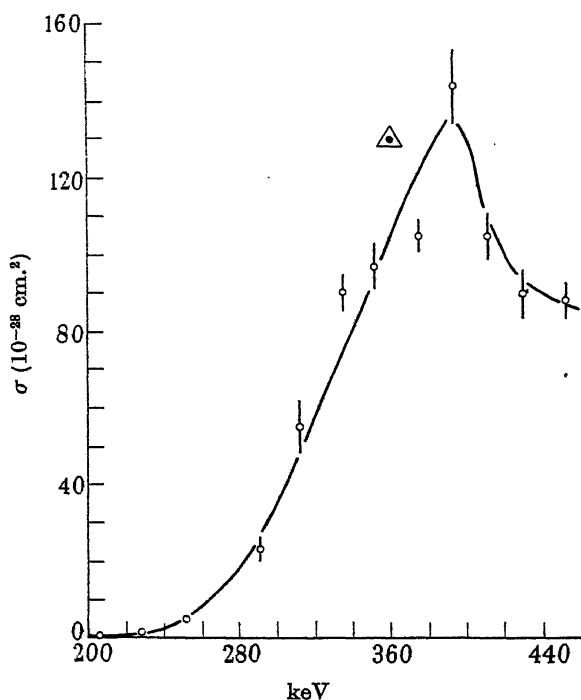
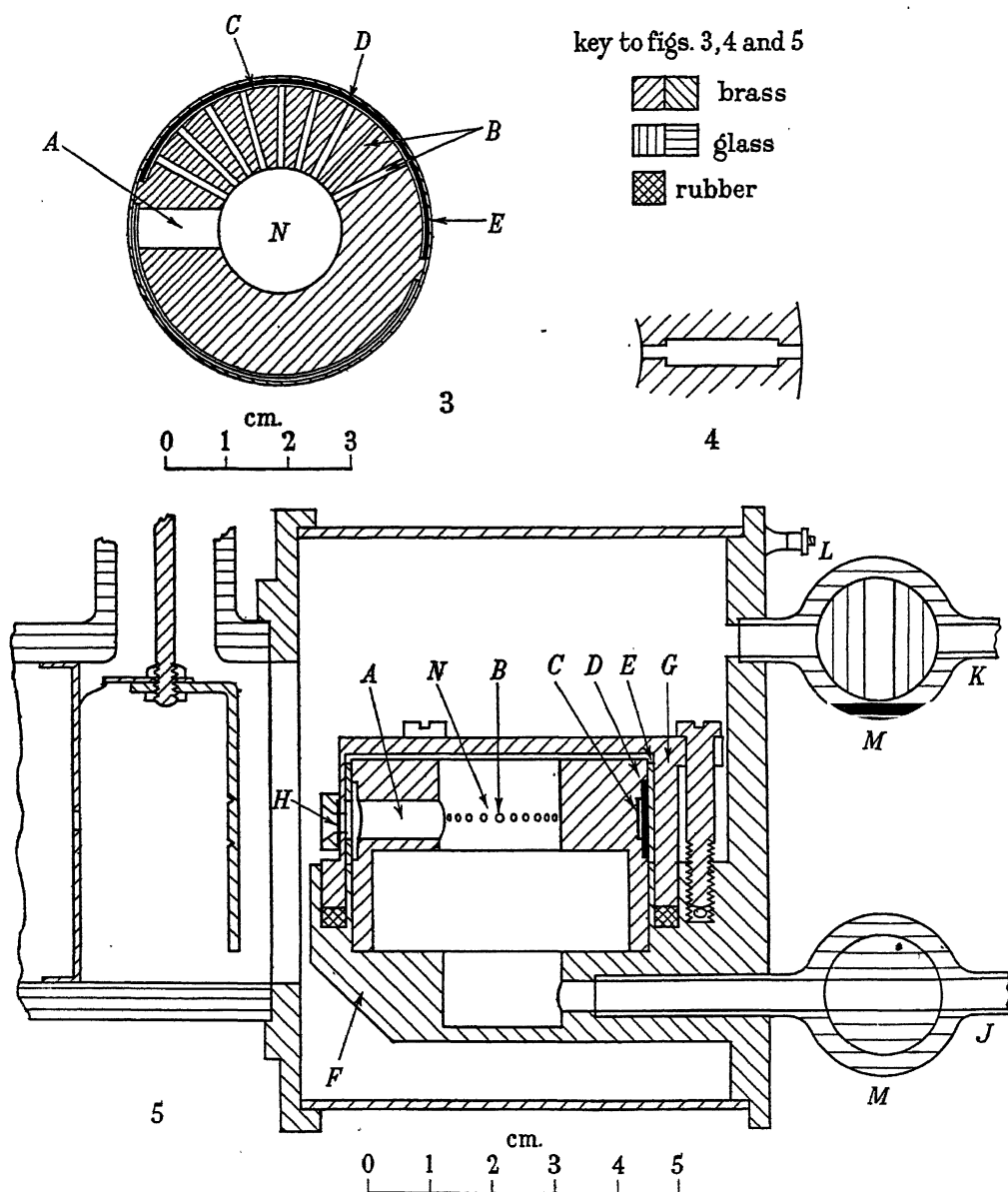


FIGURE 2. Cross-section for $\text{N}^{15}(p, \alpha)\text{C}^{12}$. Δ Observation of Holloway & Bethe (1940).

4. APPARATUS FOR MEASUREMENT OF ANGULAR DISTRIBUTION

For measurement of the angular distribution of the α -particles emitted in the reaction, a comparatively simple photographic method was adopted. A gas target of enriched ammonia (as in § 2) was used, and the particles emitted at different angles to the proton beam were recorded by allowing them to fall normally on to a photographic film bent into an arc of a circle with the target as centre. The film was Ilford Nuclear Emulsion (Type C2) and its use was suggested by Mr S. Lattimore. A plan of the target chamber and camera is shown in figure 3, the actual shape of the channels B being as shown in figure 4. The mounting of the camera on the end of the acceleration tube and the connexions to the ammonia gas supply are shown in side elevation in figure 5.



Description of parts in figures 3, 4 and 5. *A*, channel for incident proton beam; *B*, channels for emitted α -particles; *C*, aluminium foil; *D*, nuclear emulsion film; *E*, cylinder retaining film in position; *F*, main target assembly block; *G*, gas-tight cover; *H*, mica window and wire gauze; *J*, tube for introduction of ammonia; *K*, tube for equalizing pressures on each side of *H* when not exposing film; *L*, terminal for microammeter reading proton current; *M*, taps; and *N*, target volume.

The collimated proton beam was directed through the mica window *H* along the axis of the channel *A*. α -particles from disintegrations occurring near the centre of the cylindrical cavity *N* travelled outwards. Those travelling near the axis of one of the channels *B* passed straight through and into the film *D* where they were stopped, forming dense tracks perpendicular to the surface of the emulsion. The shape of the channels *B* (figure 4) was designed to trap the majority of particles which were reflected from the walls of the channels. The angular spread of the particles recorded was $\pm 3^\circ$. The channels were made at 15° intervals, the first making an angle of 30° with the incident beam and the last 150° . Scattered protons were stopped by a sufficiently thick aluminium foil *C* placed between the camera block and the film. This also served to prevent light from reaching the film. The cylinder *E*, which held the film in position, had three holes in it. One coincided with the channel *A* and a plug was used to ensure that this was so. The other two holes in *E* were small and positioned one at each end of the film. Indentations could thus be made in the film to serve as locating marks for the circular areas containing α -particle tracks.

Great care was taken to ensure that the proton beam would travel parallel to the axis of the channel *A*. As a check throughout the experiment, a willemite screen was placed round the mica window, and the beam was cut off and re-centred whenever any strong fluorescence showed. At the end of the experiments a scorch mark on the further side of the chamber, directly opposite the window, showed that the beam had travelled straight through the centre of the chamber with little divergence.

The required time of exposure was estimated from preliminary experiments and from the yields measured in § 3. At 342 keV proton energy and with a gas pressure in the target of 1.75 cm. of mercury, the exposure given was 4.5 hr. at $1.3 \mu\text{A}$ and yielded 600 tracks in the 90° area. After processing, the film was mounted with cello tape on a glass plate, ready for examination under the microscope.

5. PROCEDURE IN DEDUCING ANGULAR DISTRIBUTION

As a preliminary step, a strip of film was placed in the camera without the aluminium foil and exposed to light. On processing, the film showed black circular areas, giving the positions of the areas to be examined and their relation to the location marks as described in § 4. A copy of this was mounted alongside the film to be examined, to facilitate the finding of the areas to be searched. The appearance of the tracks was determined in another preliminary experiment by exposing a piece of film to a polonium source of α -particles at an appropriate distance.

The areas of the film which contained tracks were searched under a high-power microscope, using an oil immersion objective giving a magnification of $\times 576$. The stage had controllable horizontal traverse in both directions, and a square reticle subdivided into sixteen squares was used to specify a field of view and facilitate counting. With the aid of the reference strip, the tracks were quickly found and the edge of each area containing them was well defined. Each area corresponded to approximately 270 reticle squares. By the use of a tapping key and telephone call-meter, tracks could be counted without raising the eyes from the microscope.

Using the special nuclear film, the tracks were easily seen but there were also a few extraneous black specks in the emulsion. α -particles were distinguished from such specks by three criteria, given here in their order of application: (a) an α -particle track should start on the surface of the emulsion; (b) it should travel straight into the emulsion without any breaks; and (c) all tracks should be the same length, about $15\ \mu$. Thus an α -particle track appeared as an almost perfectly round black spot, sharply in focus at the surface and showing, generally, a small unfocused tail. Figure 6 shows a typical field of view. As the microscope fine adjustment was racked down, the spot appeared to travel along the tail. A few tracks of correct range were found which started at the surface of the emulsion but made large angles to the perpendicular, some being almost parallel to the surface. These were probably due to large-angle scattering in the emulsion and were included in the total. In any case, their number was negligible compared with the statistical error.

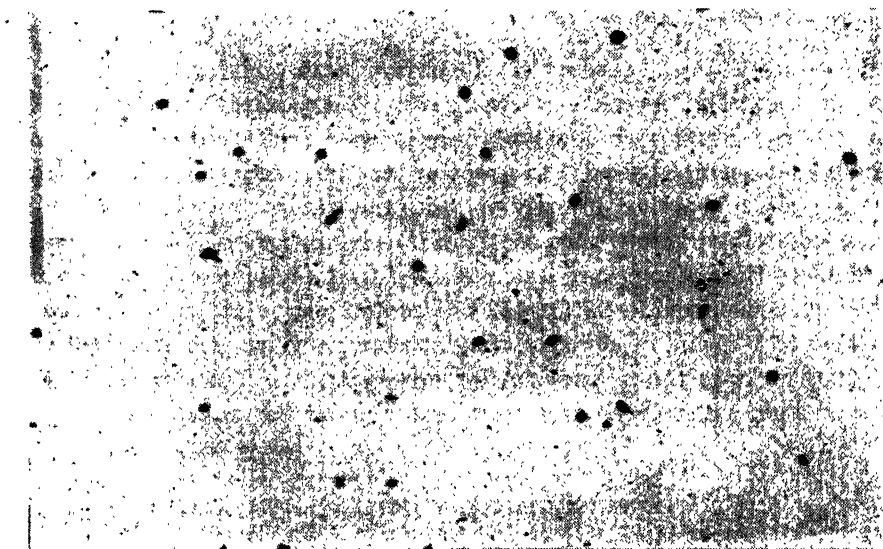


FIGURE 6. Appearance of α -particle tracks. (Magn. $\times 920$.)

When counting the number of tracks in an area, care had to be taken to avoid having gaps or overlaps in moving from one field of view to the next. From repeated counts on the same area it was estimated that this error was 1 %, and was less than the statistical error involved in the finite number of counts. On the last film studied, no attempt was made to have successive fields of view in juxtaposition and, instead, counts were recorded in a series of randomly chosen fields of view over an area. This sampling method was justified by verifying that the tracks were, in fact, randomly distributed over an area; that is, by verifying that sample counts over an area fitted well with a Poisson distribution.

Since the effective target volume varied with the angle at which the α -particles were observed, the counts for each area had to be multiplied by the appropriate

value of $\sin \theta$, where θ is the angle with the incident proton beam. A further small correction was necessary to give the distribution relative to the centre of mass of the interacting particles.

6. RESULTS FOR ANGULAR DISTRIBUTION

The investigation of the angular distribution of the α -particles was made primarily to ensure that the cross-section had been correctly calculated from the ionization chamber observations in the 90° direction. Preliminary film exposures indicated that the assumed isotropic distribution was justified, to the order of accuracy involved. However, it was thought worth while to make two more careful observations by the film method, since the angular distribution is itself of interest. These results are shown in figure 7 for 342 and 412 keV proton energy, these being the energies after passing through the mica window. The abscissa in figure 7 is $\cos \theta_c$, where θ_c is the angle between the incident and emerging particles in the centre-of-mass system. The ordinate is N_{θ_c}/N_{90} , the ratio of the numbers of tracks at θ_c and 90° . Since no observation was actually made at 90° in the centre-of-mass system, the value of N_{90} was first found by interpolation.

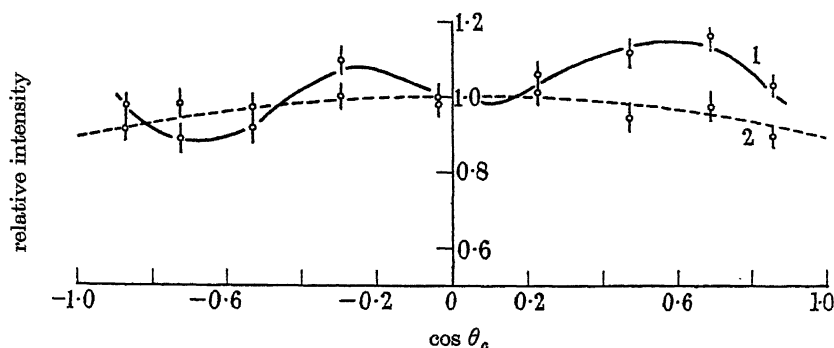


FIGURE 7. Angular distribution for 342 keV, curve 1, and for 412 keV, curve 2.

7. DISCUSSION

The graph showing the variation of the cross-section of the reaction with the energy of the incident protons (figure 2) is notable for the rather steep rise up to about 400 keV, and the indication of a maximum or resonance at that energy. The rise is much faster than would be accounted for solely by the change in the Gamow penetrability factor for protons entering an N^{15} nucleus. Observations at higher energies would be desirable but, in the absence of these, it has been assumed that a resonance exists, and its consequences have been studied. The large width of the resonance cannot be due to voltage fluctuations (~ 10 kV, at most), to straggling caused by the mica window (~ 3 kV), or to slowing down in the target (~ 8 kV). Assuming a single level of the intermediate O^{16} excited nucleus, the well-known Breit-Wigner theory gives the cross-section as

$$\sigma = \pi \lambda^2 \frac{(2J+1)}{(2s+1)(2i+1)} \frac{\Gamma_p \Gamma_\alpha}{(E - E_r)^2 + (\Gamma^2/4)},$$

where the notation of Bethe (1937) is used. The observed values of σ then give $\Gamma \doteq \Gamma_\alpha = 150$ keV, $\Gamma_p = 250$ eV, for the respective α -particle and proton widths of the resonance level of the excited O^{16} nucleus. The proton width 'without barrier' is 140 keV at the resonance. In making these deductions it is assumed that $J = 1$, $s = i = \frac{1}{2}$ (see below).

In view of Bethe's (1939) suggestion that this reaction occurs in the sun, it is of importance to utilize the data on the cross-section σ to evaluate the rate at which N^{15} is consumed. The probability per second that a given N^{15} nucleus suffers a thermonuclear reaction with H^1 will depend on the variation with the proton energy of the cross-section σ , on the temperature T , the total density ρ and the percentage of hydrogen x . The complete theory has been given for the general case by Gamow & Teller (1938), and, in applying it to the sun's core, it is usual to assume $T = 2 \times 10^7$ °C, $\rho = 80$ and $x = 35$ %. Taking, then, the present results for σ and extrapolating to low energies with the one-level Breit-Wigner formula, a lifetime of about 3×10^4 years is obtained for N^{15} in the sun. Astrophysical data on the abundance of N^{15} in the sun are not available, but it may be presumed to be not much different from the abundance on the earth, that is, about 1 : 250 compared with N^{14} . If the N^{14} and N^{15} are taking part in a cycle of reactions as suggested by Bethe, their lifetimes would be expected to be in the same ratio as their abundances. Thus the present figure of 3×10^4 years for N^{15} would correspond to 7.5×10^6 years for N^{14} . This agrees within a factor of two with Bethe's (1940) estimate of 4×10^6 years for N^{14} , based on results from the $N^{14}(p, \gamma)$ reaction.

TABLE 2. SCHEME OF ANGULAR MOMENTA AND PARITIES

nucleus	H^1	N^{15}	O^{16*}	C^{12}	He^4
nuclear spin	1/2	1/2	—	0	0
motional angular momentum	0 (<i>s</i> -wave)		—	1 (<i>p</i> -wave)	
total J	1		1	1	
nuclear parity	even	odd	odd	even	even
motional parity	even		—	odd	

The high yield of the reaction at 400 keV and the approximate isotropic distribution of the emitted α -particles are consistent with the supposition that head-on collisions between H^1 and N^{15} nuclei are mainly responsible for the disintegration in the neighbourhood of this voltage. Taking generally accepted values for the parities and spins of H^1 , N^{15} , C^{12} and He^4 , the scheme shown in table 2 would then apply. Although the motion of the product nuclei is described by a *p*-wave with one unit of angular momentum, no directions in space are specified by the spins or motional angular momentum of the initial nuclei. Consequently the angular distribution of the emitted nuclei is isotropic. While it is the most probable, the scheme of table 2 is not uniquely determined, especially as the parity of N^{15} is not well established.

Taking account of the standard deviations of the observations as shown in figure 7, curves 1 and 2, there is evidence of a slight deviation from an exactly isotropic distribution. The curve drawn through the points in curve 2 corresponds to a distribution $1 - 0.11 \cos^2 \theta_c$. The significance, or otherwise, of this result cannot be determined until other resonance levels have been established and investigated, at higher energies.

It is felt desirable to include in this discussion some remarks on experimental technique. The advantages of the photographic method for studying nuclear scattering at various angles have been stated by Chadwick, May, Pickavance & Powell (1944). In the present investigation of a nuclear disintegration (as distinct from scattering) these advantages compared with electrical counting were even more pronounced. Successive electrical counts at various angles, as well as requiring careful monitoring, would have involved an excessively long running time for the high-tension apparatus. The photographic method enabled the whole range of angles to be recorded at once. To complete the comparison with electrical counting, the extra time and trouble involved in looking at the tracks under the microscope have to be considered. In the method of Chadwick *et al.*, a flat photographic plate is placed parallel to the incident beam and a little distance from the target; the particles scattered through any particular angle lie on a cone about the incident beam and this cone intersects the plate in a hyperbola. Thus the flat-plate method involves searching the plate along a series of different hyperbolae, and necessitates a complicated setting up of the film on the microscope stage. The curved film method with discrete areas of normal tracks, as described in the present work, requires relatively simple manipulation. A thousand tracks were counted per area in half an hour, so that a whole film could be investigated in eight or ten hours. Since the intensity of disintegration products is not expected to vary rapidly with angle, the curved film method, with discrete areas, appears to satisfy all the requirements in a simple angular distribution experiment. However, measurements involving the detection of particles of various ranges, or the variation of range with angle, require the tracks to be nearly parallel to the surface of the emulsion as in the methods of Chadwick *et al.* (1944) or Rubin (1947); alternatively—at high yields—a cloud chamber will give the information.

The authors are much indebted to Sir George Thomson for advice and discussion throughout the investigation and for practical help in the early stages; also to Mr L. Norriss for maintaining and operating the equipment. The major part of the apparatus was made to the order of the Central Research Fund Committee of the University of London and was granted to the authors on an extended loan.

REFERENCES

- Bethe, H. A. 1937 *Rev. Mod. Phys.* **9**, 69.
Bethe, H. A. 1939 *Phys. Rev.* **55**, 434.
Bethe, H. A. 1940 *Astrophys. J.* **92**, 118.
Burcham, W. E. & Smith, C. L. 1939 *Nature*, **143**, 795.
Chadwick, J., May, A. N., Pickavance, T. G. & Powell, C. F. 1944 *Proc. Roy. Soc. A*, **183**, 1.

- Craggs, J. D. 1942 *Proc. Phys. Soc.* **54**, 439.
 Gamow, G. & Teller, E. 1938 *Phys. Rev.* **53**, 608.
 Holloway, M. G. & Bethe, H. A. 1940 *Phys. Rev.* **57**, 747.
 Livingston, M. S. & Bethe, H. A. 1937 *Rev. Mod. Phys.* **9**, 270.
 Rubin, S. 1947 *Phys. Rev.* **72**, 1176.
 Smee, J. F. 1944 *J. Instn Elect. Engrs*, **91**, Part 1, 422.
 Wynn-Williams, C. E. 1936 *Progr. Rep. Phys. Soc.* **3**, 239.

Unsteady compressible flow in two dimensions

BY T. V. DAVIES, *King's College, London*

(Communicated by G. Temple, F.R.S.—Received 27 May 1949)

The paper presents an alternative to the well-known linear perturbation theory of two-dimensional compressible flow by superposing a small perturbation upon the correct basic irrotational flow pattern. When the latter is known the perturbation field is determined by partial differentiation of a function which satisfies a linear partial differential equation of the second order. A particular example of the theory is investigated in which the basic flow is circular and bounded internally by an oscillating circular cylinder, and the author restricts himself to the problem of wave propagation.

INTRODUCTION

The linear perturbation theory for unsteady motion in two dimensions, in which a small unsteady perturbation is superposed upon a steady rectilinear basic flow, has very obvious limitations, and the present paper is an attempt to improve upon this theory by substituting for the linear basic flow the correct irrotational flow pattern. This idea is not new in incompressible theory (see Lewis 1938), but the application here is to compressible fluids.

The flow is considered to be composed of a steady basic flow, assumed to be compressible but irrotational, together with an unsteady perturbation flow which is small compared with the basic flow. The basic flow will be denoted, as far as is practicable, in the usual notation; at a point in the physical plane c will denote the basic velocity and θ its direction (relative to x -axis). It is found convenient to express the perturbation components of velocity along and perpendicular to c ; when this is done it is possible to express these perturbation components and the unsteady pressure in terms of a function α and its partial derivatives. The function α is shown to satisfy a second-order, linear partial differential equation. To illustrate the theory the particular example of unsteady flow around an oscillating cylinder is discussed, and in this problem we shall be interested principally in wave propagation.

1. EQUATIONS OF MOTION IN TWO DIMENSIONS

If dashes are used to represent the general expressions for velocity components and pressure, the equations of motion may be written

$$\left. \begin{aligned} \frac{\partial u'}{\partial t} + u' \frac{\partial u'}{\partial x} + v' \frac{\partial u'}{\partial y} &= -\frac{1}{\rho'} \frac{\partial p'}{\partial x} - \frac{\partial \Omega}{\partial x}, \\ \frac{\partial v'}{\partial t} + u' \frac{\partial v'}{\partial x} + v' \frac{\partial v'}{\partial y} &= -\frac{1}{\rho'} \frac{\partial p'}{\partial y} - \frac{\partial \Omega}{\partial y} \end{aligned} \right\} \quad (1.1)$$

Introducing the vorticity $\zeta' = \frac{\partial v'}{\partial x} - \frac{\partial u'}{\partial y}$, (1.2)

and assuming that the pressure is a function of the density, equations (1.1) become

$$\begin{aligned} \frac{\partial u'}{\partial t} - v'\zeta' &= -\frac{\partial}{\partial x} \left\{ \int \frac{dp'}{\rho'} + \frac{1}{2}(u'^2 + v'^2) + \Omega \right\}, \\ \frac{\partial v'}{\partial t} + u'\zeta' &= -\frac{\partial}{\partial y} \left\{ \int \frac{dp'}{\rho'} + \frac{1}{2}(u'^2 + v'^2) + \Omega \right\}. \end{aligned} \quad (1.3)$$

If we write $\chi' = \int \frac{dp'}{\rho'} + \frac{1}{2}(u'^2 + v'^2) + \Omega$, (1.4)

the set (1.3) and the equation of continuity are

$$\frac{\partial u'}{\partial t} - v'\zeta' = -\frac{\partial \chi'}{\partial x}, \quad \frac{\partial v'}{\partial t} + u'\zeta' = -\frac{\partial \chi'}{\partial y}, \quad (1.5)$$

$$\frac{\partial \rho'}{\partial t} + \frac{\partial}{\partial x}(\rho'u') + \frac{\partial}{\partial y}(\rho'v') = 0. \quad (1.6)$$

2. THE BASIC FLOW

Denoting the basic flow by capitals, we now assume that

$$u' = U + u, \quad v' = V + v, \quad \rho' = P + \rho, \quad p' = \Pi + p, \quad \chi' = \chi_0 + \chi_1. \quad (2.1)$$

The irrotational condition of the basic flow is

$$\frac{\partial V}{\partial x} = \frac{\partial U}{\partial y}. \quad (2.2)$$

The density P being independent of the time, the equation of continuity (1.6) leads to

$$\frac{\partial}{\partial x}(PU) + \frac{\partial}{\partial y}(PV) = 0, \quad (2.3)$$

and with these assumptions, equations (1.5) become

$$0 = \frac{\partial \chi_0}{\partial x} = \frac{\partial \chi_0}{\partial y},$$

so that $\int \frac{d\Pi}{P} + \frac{1}{2}(U^2 + V^2) + \Omega = \text{an absolute constant} = \chi_0$. (2.4)

We write $c^2 = U^2 + V^2$, (2.5)

and hence $\int \frac{d\Pi}{P} + \frac{1}{2}c^2 + \Omega = \chi_0$. (2.6)

Equations (2.2) and (2.3) show that two functions ϕ and ψ may be introduced conveniently, of the nature of velocity potential and stream functions respectively. We take

$$U = \frac{\partial \phi}{\partial x}, \quad V = \frac{\partial \phi}{\partial y}, \quad PU = P_0 \frac{\partial \psi}{\partial y}, \quad PV = -P_0 \frac{\partial \psi}{\partial x}, \quad (2.7)$$

following the usual conventions in notation. P_0 will represent in (2.7) the density at a stagnation point.

3. THE PERTURBATION FLOW AND ITS TRANSFORMATION

Substituting the expressions (2.1) into the general equations of motion (1.5) and using the results of § 2, the equations governing the perturbation flow are

$$\frac{\partial u}{\partial t} - V\zeta = -\frac{\partial \chi_1}{\partial x}, \quad \frac{\partial v}{\partial t} + U\zeta = -\frac{\partial \chi_1}{\partial y}, \quad (3.1)$$

where

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}, \quad (3.2)$$

and the unsteady motion is assumed to be sufficiently small to neglect the products ζu and ζv .

In place of the components (u, v) we now introduce the perturbation velocity components (u_1, v_1) , the former being along the direction of c , the latter perpendicular to c (θ increasing). Thus we have (see figure 1)

$$u = u_1 \cos \theta - v_1 \sin \theta, \quad v = u_1 \sin \theta + v_1 \cos \theta. \quad (3.3)$$

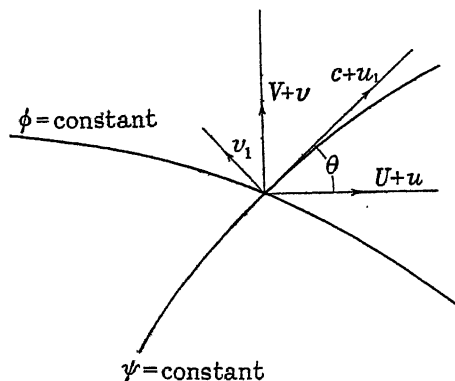


FIGURE 1

We introduce also new independent variables (ϕ, ψ) in place of (x, y) the transformation of co-ordinates being effected by the relations (see (2.7))

$$P \frac{\partial \phi}{\partial x} = P_0 \frac{\partial \psi}{\partial y}, \quad P \frac{\partial \phi}{\partial y} = -P_0 \frac{\partial \psi}{\partial x}. \quad (3.4)$$

Using these relations we have

$$\left. \begin{aligned} \frac{\partial}{\partial x} &= \frac{\partial \phi}{\partial x} \frac{\partial}{\partial \phi} + \frac{\partial \psi}{\partial x} \frac{\partial}{\partial \psi} = U \frac{\partial}{\partial \phi} - \frac{PV}{P_0} \frac{\partial}{\partial \psi}, \\ \frac{\partial}{\partial y} &= \frac{\partial \phi}{\partial y} \frac{\partial}{\partial \phi} + \frac{\partial \psi}{\partial y} \frac{\partial}{\partial \psi} = V \frac{\partial}{\partial \phi} + \frac{PU}{P_0} \frac{\partial}{\partial \psi}, \end{aligned} \right\} \quad (3.5)$$

and furthermore

$$U = c \cos \theta, \quad V = c \sin \theta. \quad (3.6)$$

The expression for the vorticity ζ in terms of $(\phi, \psi, c, \theta, u_1, v_1)$ is as follows:

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \left(U \frac{\partial}{\partial \phi} - \frac{PV}{P_0} \frac{\partial}{\partial \psi} \right) \{u_1 \sin \theta + v_1 \cos \theta\} - \left(V \frac{\partial}{\partial \phi} + \frac{PU}{P_0} \frac{\partial}{\partial \psi} \right) \{u_1 \cos \theta - v_1 \sin \theta\}.$$

Thus
$$\zeta = c \left\{ u_1 \frac{\partial \theta}{\partial \phi} + \frac{\partial v_1}{\partial \phi} - \frac{P}{P_0} \frac{\partial u_1}{\partial \psi} + \frac{P v_1}{P_0} \frac{\partial \theta}{\partial \psi} \right\}, \quad (3.7)$$

and the equations of motion (3.1) now become

$$\begin{aligned} \cos \theta \frac{\partial u_1}{\partial t} - \sin \theta \frac{\partial v_1}{\partial t} - c^2 \sin \theta \left\{ u_1 \frac{\partial \theta}{\partial \phi} + \frac{\partial v_1}{\partial \phi} - \frac{P}{P_0} \frac{\partial u_1}{\partial \psi} + \frac{P v_1}{P_0} \frac{\partial \theta}{\partial \psi} \right\} \\ = -c \cos \theta \frac{\partial \chi_1}{\partial \phi} + \frac{P c \sin \theta}{P_0} \frac{\partial \chi_1}{\partial \psi}, \\ \sin \theta \frac{\partial u_1}{\partial t} + \cos \theta \frac{\partial v_1}{\partial t} + c^2 \cos \theta \left\{ u_1 \frac{\partial \theta}{\partial \phi} + \frac{\partial v_1}{\partial \phi} - \frac{P}{P_0} \frac{\partial u_1}{\partial \psi} + \frac{P v_1}{P_0} \frac{\partial \theta}{\partial \psi} \right\} \\ = -c \sin \theta \frac{\partial \chi_1}{\partial \phi} - \frac{P c \cos \theta}{P_0} \frac{\partial \chi_1}{\partial \psi}. \end{aligned}$$

These equations lead to
$$\frac{\partial u_1}{\partial t} = -c \frac{\partial \chi_1}{\partial \phi}, \quad (3.8)$$

$$\frac{\partial v_1}{\partial t} + c^2 \left\{ u_1 \frac{\partial \theta}{\partial \phi} + \frac{\partial v_1}{\partial \phi} - \frac{P}{P_0} \frac{\partial u_1}{\partial \psi} + \frac{P v_1}{P_0} \frac{\partial \theta}{\partial \psi} \right\} = -\frac{P c}{P_0} \frac{\partial \chi_1}{\partial \psi}. \quad (3.9)$$

Since c is independent of t equation (3.8) shows that we can introduce a function α such that

$$u_1 = c \frac{\partial \alpha}{\partial \phi}, \quad (3.10)$$

$$\chi_1 = -\frac{\partial \alpha}{\partial t}. \quad (3.11)$$

4. THE SOLUTION FOR v_1 IN TERMS OF α , ϕ , ETC.

Before finding the form of solution for v_1 we develop from (2.7), (3.4) and (3.6) a more convenient set of relations expressing the derivatives of c with respect to ϕ and ψ in terms of the derivatives of θ .

From (2.7) if we write $z = x + iy$,

$$\begin{aligned} c e^{-i\theta} dz &= c(\cos \theta - i \sin \theta)(dx + i dy) \\ &= (U - iV)(dx + i dy) \\ &= U dx + V dy - i(V dx - U dy) \\ &= \frac{\partial \phi}{\partial x} dx + \frac{\partial \phi}{\partial y} dy - i \left(-\frac{P_0}{P} \frac{\partial \psi}{\partial x} dx - \frac{P_0}{P} \frac{\partial \psi}{\partial y} dy \right) \\ &= d\phi + i \frac{P_0}{P} d\psi, \end{aligned}$$

hence
$$dz = \frac{1}{c} e^{i\theta} \left\{ d\phi + i \frac{P_0}{P} d\psi \right\}, \quad (4.1)$$

and thus
$$\frac{\partial z}{\partial \phi} = \frac{e^{i\theta}}{c}, \quad \frac{\partial z}{\partial \psi} = i \frac{P_0}{P} \frac{e^{i\theta}}{c}.$$

It follows that

$$\frac{\partial}{\partial \psi} \left(\frac{e^{i\theta}}{c} \right) = \frac{\partial}{\partial \phi} \left(i \frac{P_0}{P} \frac{e^{i\theta}}{c} \right),$$

$$i \frac{\partial \theta}{c \partial \psi} - \frac{1}{c^2} \frac{\partial c}{\partial \psi} = i \frac{\partial}{\partial \phi} \left(\frac{P_0}{Pc} \right) - \frac{P_0}{Pc} \frac{\partial \theta}{\partial \phi}.$$

Equating real and imaginary parts

$$\frac{1}{c} \frac{\partial \theta}{\partial \psi} = \frac{\partial}{\partial \phi} \left(\frac{P_0}{Pc} \right) = - \frac{P_0}{P^2 c^2} \frac{\partial}{\partial \phi} (Pc), \quad \frac{1}{c^2} \frac{\partial c}{\partial \psi} = \frac{P_0}{Pc} \frac{\partial \theta}{\partial \phi}. \quad (4.2)$$

These relations will be required in the form

$$\frac{\partial}{\partial \phi} (Pc) + \frac{P^2 c}{P_0} \frac{\partial \theta}{\partial \psi} = 0, \quad P \frac{\partial c}{\partial \psi} - P_0 c \frac{\partial \theta}{\partial \phi} = 0. \quad (4.3)$$

These are the reciprocals of the Tschiapugin equations (see Carrier & Carlson 1946). We can now consider the expression (3.9). The form of the solution for u_1 suggests the form

$$v_1 = + \frac{Pc}{P_0} \frac{\partial \alpha}{\partial \psi}, \quad (4.4)$$

for the v_1 perturbation. If we substitute for u_1 , v_1 , χ_1 in (3.9) we obtain

$$0 = + \frac{Pc}{P_0} \frac{\partial^2 \alpha}{\partial t \partial \psi} + c^2 \left\{ c \frac{\partial \alpha}{\partial \phi} \frac{\partial \theta}{\partial \phi} + \frac{\partial}{\partial \phi} \left[\frac{Pc}{P_0} \frac{\partial \alpha}{\partial \psi} \right] - \frac{P}{P_0} \frac{\partial}{\partial \psi} \left(c \frac{\partial \alpha}{\partial \phi} \right) + \frac{P}{P_0} \frac{\partial \theta}{\partial \psi} \frac{Pc}{P_0} \frac{\partial \alpha}{\partial \psi} \right\} - \frac{Pc}{P_0} \frac{\partial^2 \alpha}{\partial \psi \partial t}$$

$$= c^2 \left\{ \frac{\partial \alpha}{\partial \phi} \left[c \frac{\partial \theta}{\partial \phi} - \frac{P}{P_0} \frac{\partial c}{\partial \psi} \right] + \frac{\partial \alpha}{\partial \psi} \left[\frac{\partial}{\partial \phi} \left(\frac{Pc}{P_0} \right) + \frac{P^2 c}{P_0^2} \frac{\partial \theta}{\partial \psi} \right] \right\}. \quad (4.5)$$

The right-hand side of (4.5) is identically zero in virtue of the relations (4.3), P_0 being a constant. It is evident then that (4.4) is a solution for v_1 , but it is not the general solution, since we may add to the right-hand side of (4.4) any solutions of the partial differential equation

$$\frac{\partial v_1}{\partial t} + c^2 \frac{\partial v_1}{\partial \phi} + \frac{c^2 P}{P_0} \frac{\partial \theta}{\partial \psi} v_1 = 0. \quad (4.6)$$

If in (4.6) we replace the coefficient of v_1 , using (4.2), the equation becomes

$$\frac{\partial v_1}{\partial t} + c^2 \frac{\partial v_1}{\partial \phi} - \frac{c}{P} \frac{\partial}{\partial \phi} (Pc) v_1 = 0,$$

and if we write

$$v_1 = \frac{Pc}{P_0} \beta, \quad (4.7)$$

then

$$\frac{Pc}{P_0} \frac{\partial \beta}{\partial t} + c^2 \frac{Pc}{P_0} \frac{\partial \beta}{\partial \phi} + c^2 \beta \frac{\partial}{\partial \phi} \left(\frac{Pc}{P_0} \right) - c^2 \beta \frac{\partial}{\partial \phi} \left(\frac{Pc}{P_0} \right) = 0,$$

that is,

$$\frac{\partial \beta}{\partial t} + c^2 \frac{\partial \beta}{\partial \phi} = 0. \quad (4.8)$$

Thus in general we may state that

$$v_1 = \frac{Pc}{P_0} \left\{ \frac{\partial \alpha}{\partial \psi} + \beta \right\}, \quad (4.9)$$

where β is any solution of the equation (4.8). It follows that the general expressions for (u_1, v_1, χ_1) are

$$u_1 = c \frac{\partial \alpha}{\partial \phi}, \quad v_1 = \frac{Pc}{P_0} \left(\frac{\partial \alpha}{\partial \psi} + \beta \right), \quad \chi_1 = -\frac{\partial \alpha}{\partial t}. \quad (4.10)$$

5. THE CONDITION FOR IRROTATIONAL FLOW AND INTERPRETATION OF β

The condition for irrotational flow is

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = 0, \quad (5.1)$$

and using the transformed expression for ζ , namely, (3.7), the condition is

$$\frac{\zeta}{c} = u_1 \frac{\partial \theta}{\partial \phi} + \frac{\partial v_1}{\partial \phi} - \frac{P}{P_0} \frac{\partial u_1}{\partial \psi} + \frac{Pv_1}{P_0} \frac{\partial \theta}{\partial \psi} = 0. \quad (5.2)$$

Substituting for u_1 and v_1 from (4.10) this becomes

$$\frac{\zeta}{c} = \frac{Pc}{P_0} \frac{\partial \beta}{\partial \phi}.$$

Thus the condition for irrotational motion is

$$\frac{\partial \beta}{\partial \phi} = 0. \quad (5.3)$$

It will be noted from (4.8) that when (5.3) is satisfied, then so also is

$$\frac{\partial \beta}{\partial t} = 0, \quad (5.4)$$

and β , in irrotational motion, will therefore be a function of ψ only, that is, β will be constant along the curves $\psi = \text{constant}$.

Thus, when the unsteady motion is a perturbation upon a main flow, the β function which occurs in (4.10) is related to the vorticity by

$$\zeta = \frac{Pc^2}{P_0} \frac{\partial \beta}{\partial \phi} = -\frac{P}{P_0} \frac{\partial \beta}{\partial t}, \quad (5.5)$$

and equation (5.5) now forms the interpretation of β . It will be noted also from (5.5) and (4.8) that the latter equation expresses the fact that vorticity is convected with the fluid (Carrier & Carlson 1946).

6. THE EXPRESSION FOR ρ IN TERMS OF α WHEN THE ADIABATIC RELATION IS ASSUMED TO HOLD IN THE COMPRESSIBLE FLUID

If in (1.4) we assume that the relation between p' and ρ' is

$$p' = \kappa \rho'^\gamma, \quad (6.1)$$

then

$$\chi' = \frac{\kappa\gamma}{\gamma-1} \rho'^{\gamma-1} + \frac{1}{2}(u'^2 + v'^2) + \Omega. \quad (6.2)$$

Assuming also that in the basic flow

$$\Pi = \kappa P^\gamma, \quad (6.3)$$

it follows that

$$\begin{aligned} \chi_0 + \chi_1 &= \frac{\kappa\gamma}{\gamma-1} (P + \rho)^{\gamma-1} + \frac{1}{2}\{(U+u)^2 + (V+v)^2\} + \Omega \\ &= \frac{\kappa\gamma}{\gamma-1} P^{\gamma-1} \left\{ 1 + (\gamma-1) \frac{\rho}{P} \right\} + \frac{1}{2}(U^2 + V^2) + Uu + Vv + \Omega, \end{aligned}$$

where all powers of the second and higher order are neglected, and using (2.6), namely,

$$\chi_0 = \frac{\kappa\gamma}{\gamma-1} P^{\gamma-1} + \frac{1}{2}c^2 + \Omega, \quad (6.4)$$

we obtain

$$\begin{aligned} \chi_1 &= \frac{\kappa\gamma}{\gamma-1} P^{\gamma-1} (\gamma-1) \frac{\rho}{P} + Uu + Vv \\ &= \frac{\gamma\Pi}{P} \frac{\rho}{P} + Uu + Vv. \end{aligned} \quad (6.5)$$

Substituting for U, V, u, v , from (3.3)

$$\begin{aligned} \chi_1 &= \frac{\gamma\Pi}{P} \frac{\rho}{P} + c \cos \theta \{u_1 \cos \theta - v_1 \sin \theta\} + c \sin \theta \{u_1 \sin \theta + v_1 \cos \theta\} \\ &= \frac{\gamma\Pi}{P} \frac{\rho}{P} + cu_1. \end{aligned}$$

Writing
$$a^2 = \frac{\gamma\Pi}{P} = \kappa\gamma P^{\gamma-1}, \quad a_0^2 = \kappa\gamma P_0^{\gamma-1}, \quad (6.6)$$

where a is the local speed of sound, this relation for ρ becomes

$$-\frac{\partial \alpha}{\partial t} = a^2 \frac{\rho}{P} + c^2 \frac{\partial \alpha}{\partial \phi},$$

hence
$$\frac{\rho}{P} = -\frac{1}{a^2} \frac{\partial \alpha}{\partial t} - \frac{c^2}{a^2} \frac{\partial \alpha}{\partial \phi}. \quad (6.7)$$

From (2.1), (6.1) and (6.3) it follows that

$$p = a^2 \rho,$$

and thus from (6.7) we obtain

$$p = -P \left(\frac{\partial \alpha}{\partial t} + c^2 \frac{\partial \alpha}{\partial \phi} \right). \quad (6.7)'$$

The normal form of the relation (6.4) is included here for future reference. When Ω is omitted from the equation it may be written

$$\frac{1}{\gamma-1} a^2 + \frac{1}{2} c^2 = \text{constant},$$

that is,
$$a^2 = a_0^2 - \frac{1}{2}(\gamma-1)c^2, \quad (6.8)$$

where a_0 is the velocity of sound at a stagnation point.

The alternative form of (6.8) is

$$\frac{a^2}{a_0^2} = 1 - \frac{(\gamma-1)c^2}{2a_0^2}. \quad (6.9)$$

The maximum velocity which c may attain is c_m , where

$$c_m^2 = \frac{2}{\gamma-1} a_0^2 \quad (= 5a_0^2 \text{ for air}), \quad (6.10)$$

and using this new constant c_m

$$\frac{a^2}{a_0^2} = 1 - \frac{c^2}{c_m^2} = 1 - \tau, \quad (6.11)$$

where τ is a non-dimensional variable, lying in the range $0 \leq \tau \leq 1$,

$$\tau = \frac{c^2}{c_m^2}. \quad (6.12)$$

From (6.6)
$$\frac{a^2}{a_0^2} = \left(\frac{P}{P_0}\right)^{\gamma-1} = 1 - \tau = 1 - \frac{c^2}{c_m^2}, \quad (6.13)$$

and if we write
$$\beta_0 = \frac{1}{\gamma-1} \quad (= 5/2 \text{ for air}), \quad (6.14)$$

then
$$\frac{P}{P_0} = \left(1 - \frac{c^2}{c_m^2}\right)^{\beta_0} = (1 - \tau)^{\beta_0}. \quad (6.15)$$

7. THE EQUATION OF CONTINUITY IN THE TRANSFORMED CO-ORDINATES

In the general equation of continuity

$$\frac{\partial \rho'}{\partial t} + \frac{\partial}{\partial x}(\rho' u') + \frac{\partial}{\partial y}(\rho' v') = 0, \quad (7.1)$$

making the substitutions (2.1) and using (2.3), this becomes

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(Pu + \rho U) + \frac{\partial}{\partial y}(Pv + \rho V) = 0, \quad (7.2)$$

where we neglect the second-order terms ρu and ρv . Equation (7.2) may be written

$$\frac{\partial \rho}{\partial t} + P \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) + u \frac{\partial P}{\partial x} + v \frac{\partial P}{\partial y} + U \frac{\partial \rho}{\partial x} + V \frac{\partial \rho}{\partial y} + \rho \left(\frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} \right) = 0. \quad (7.3)$$

We take pairs of terms in (7.3) and transform to (ϕ, ψ) co-ordinates using the relations of § 3:

$$\begin{aligned} \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} &= \left(U \frac{\partial}{\partial \phi} - \frac{PV}{P_0} \frac{\partial}{\partial \psi} \right) (u_1 \cos \theta - v_1 \sin \theta) + \left(V \frac{\partial}{\partial \phi} + \frac{PU}{P_0} \frac{\partial}{\partial \psi} \right) (u_1 \sin \theta + v_1 \cos \theta) \\ &= c \frac{\partial u_1}{\partial \phi} - \frac{\partial \theta}{\partial \phi} + \frac{Pcu_1}{P_0} \frac{\partial \theta}{\partial \psi} + \frac{Pc \partial v_1}{P_0 \partial \psi}, \end{aligned} \quad (7.4)$$

$$\begin{aligned}
 u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} &= (u_1 \cos \theta - v_1 \sin \theta) \left\{ c \cos \theta \frac{\partial}{\partial \phi} - \frac{Pc \sin \theta}{P_0} \frac{\partial}{\partial \psi} \right\} \\
 &\quad + (u_1 \sin \theta + v_1 \cos \theta) \left\{ c \sin \theta \frac{\partial}{\partial \phi} + \frac{Pc \cos \theta}{P_0} \frac{\partial}{\partial \psi} \right\} \\
 &= u_1 c \frac{\partial}{\partial \phi} + \frac{v_1 Pc}{P_0} \frac{\partial}{\partial \psi},
 \end{aligned} \tag{7.5}$$

$$\begin{aligned}
 U \frac{\partial}{\partial x} + V \frac{\partial}{\partial y} &= c \cos \theta \left\{ c \cos \theta \frac{\partial}{\partial \phi} - \frac{Pc \sin \theta}{P_0} \frac{\partial}{\partial \psi} \right\} + c \sin \theta \left\{ c \sin \theta \frac{\partial}{\partial \phi} + \frac{Pc \cos \theta}{P_0} \frac{\partial}{\partial \psi} \right\} \\
 &= c^2 \frac{\partial}{\partial \phi},
 \end{aligned} \tag{7.6}$$

$$\begin{aligned}
 \frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} &= \left(c \cos \theta \frac{\partial}{\partial \phi} - \frac{Pc \sin \theta}{P_0} \frac{\partial}{\partial \psi} \right) (c \cos \theta) + \left(c \sin \theta \frac{\partial}{\partial \phi} + \frac{Pc \cos \theta}{P_0} \frac{\partial}{\partial \psi} \right) (c \sin \theta) \\
 &= c \left\{ \frac{\partial c}{\partial \phi} + \frac{Pc}{P_0} \frac{\partial \theta}{\partial \psi} \right\}.
 \end{aligned} \tag{7.7}$$

Using (4.3) the relation (7.7) simplifies to

$$\begin{aligned}
 \frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} &= \frac{c}{P} \left\{ P \frac{\partial c}{\partial \phi} + \frac{P^2 c}{P_0} \frac{\partial \theta}{\partial \psi} \right\} \\
 &= \frac{c}{P} \left\{ -c \frac{\partial P}{\partial \phi} \right\} = -\frac{c^2 \partial P}{P \partial \phi}.
 \end{aligned} \tag{7.8}$$

Equation (7.3) now becomes

$$\frac{\partial \rho}{\partial t} + P \left\{ c \frac{\partial u_1}{\partial \phi} - c v_1 \frac{\partial \theta}{\partial \phi} + \frac{Pc u_1}{P_0} \frac{\partial \theta}{\partial \psi} + \frac{Pc \partial v_1}{P_0 \partial \psi} \right\} + u_1 c \frac{\partial P}{\partial \phi} + \frac{v_1 Pc}{P_0} \frac{\partial P}{\partial \psi} + c^2 \frac{\partial \rho}{\partial \phi} - \frac{c^2 \rho \partial P}{P \partial \phi} = 0,$$

hence the equation of continuity in the transformed system is

$$\frac{\partial \rho}{\partial t} + Pc^2 \frac{\partial}{\partial \phi} \left(\frac{u_1}{c} \right) + \frac{Pc^2}{P_0} \left\{ \frac{P \partial v_1}{c \partial \psi} + v_1 \frac{\partial}{\partial \psi} \left(\frac{P}{c} \right) \right\} + Pc^2 \frac{\partial}{\partial \phi} \left(\frac{\rho}{P} \right) = 0,$$

$$\text{that is,} \quad \frac{\partial}{\partial t} \left(\frac{\rho}{c^2 P} \right) + \frac{\partial}{\partial \phi} \left(\frac{u_1}{c} \right) + \frac{1}{P_0} \frac{\partial}{\partial \psi} \left(\frac{v_1 P}{c} \right) + \frac{\partial}{\partial \phi} \left(\frac{\rho}{P} \right) = 0. \tag{7.9}$$

When we substitute for ρ , u_1 , v_1 in this equation we obtain

$$-\frac{\partial}{\partial t} \left\{ \frac{P}{c^2 P} \left[\frac{1}{a^2} \frac{\partial \alpha}{\partial t} + \frac{c^2 \partial \alpha}{a^2 \partial \phi} \right] \right\} + \frac{\partial}{\partial \phi} \left(\frac{\partial \alpha}{\partial \phi} \right) + \frac{1}{P_0} \frac{\partial}{\partial \psi} \left\{ \frac{P^2}{P_0} \left(\frac{\partial \alpha}{\partial \psi} + \beta \right) \right\} + \frac{\partial}{\partial \phi} \left\{ -\frac{1}{a^2} \frac{\partial \alpha}{\partial t} - \frac{c^2 \partial \alpha}{a^2 \partial \phi} \right\} = 0,$$

that is,

$$-\frac{1}{a^2 c^2} \frac{\partial^2 \alpha}{\partial t^2} - \frac{2}{a^2} \frac{\partial^2 \alpha}{\partial t \partial \phi} + \frac{\partial^2 \alpha}{\partial \phi^2} + \frac{\partial}{\partial \psi} \left\{ \frac{P^2}{P_0^2} \left(\frac{\partial \alpha}{\partial \psi} + \beta \right) \right\} - \frac{\partial \alpha}{\partial t} \frac{\partial}{\partial \phi} \left(\frac{1}{a^2} \right) - \frac{\partial}{\partial \phi} \left(\frac{c^2 \partial \alpha}{a^2 \partial \phi} \right) = 0,$$

$$\text{or} \quad -\frac{1}{a^2 c^2} \frac{\partial^2 \alpha}{\partial t^2} - \frac{2}{a^2} \frac{\partial^2 \alpha}{\partial t \partial \phi} - \frac{\partial \alpha}{\partial t} \frac{\partial}{\partial \phi} \left(\frac{1}{a^2} \right) + \frac{\partial}{\partial \phi} \left\{ \left(1 - \frac{c^2}{a^2} \right) \frac{\partial \alpha}{\partial \phi} \right\} + \frac{\partial}{\partial \psi} \left\{ \frac{P^2}{P_0^2} \left(\frac{\partial \alpha}{\partial \psi} + \beta \right) \right\} = 0. \tag{7.10}$$

If required the independent variables in this equation may be transformed from (ϕ, ψ) to (c, θ) . In the particular example to follow this transformation is easily effected.

It is of interest to state here the simplification which takes place in (7.10) when the fluid is incompressible.

If we write $P = P_0$ in (4.10) the expressions for the velocity components become

$$u_1 = c \frac{\partial \alpha}{\partial \phi}, \quad v_1 = c \left(\frac{\partial \alpha}{\partial \psi} + \beta \right), \quad \chi_1 = - \frac{\partial \alpha}{\partial t}, \quad (7.11)$$

and taking $P = P_0$ and $\rho = 0$ in (7.9), the equation of continuity becomes

$$\frac{\partial}{\partial \phi} \left(\frac{u_1}{c} \right) + \frac{\partial}{\partial \psi} \left(\frac{v_1}{c} \right) = 0. \quad (7.12)$$

Substituting for u_1 and v_1 it follows that the partial differential equation for α is

$$\frac{\partial^2 \alpha}{\partial \phi^2} + \frac{\partial^2 \alpha}{\partial \psi^2} + \frac{\partial \beta}{\partial \psi} = 0. \quad (7.13)$$

The development of this equation is left to a future paper.

8. SPECIAL CASE WHEN THE BASIC FLOW IS CIRCULAR AND BOUNDED INTERNALLY BY AN OSCILLATING CYLINDER

We can derive the standard results for circular flow from (4.2). If we assume

$$\phi = Rc_m \theta, \quad (8.1)$$

where c_m is the maximum fluid velocity, and R is a constant of dimensions length, then from (4.3), since $\frac{\partial \theta}{\partial \psi} = 0$, we have $\frac{\partial}{\partial \phi} (Pc) = 0$. Thus Pc is dependent upon ψ only, and the relation between these quantities is

$$\frac{P}{P_0} \frac{\partial c}{\partial \psi} = \frac{1}{R} \frac{c}{c_m}, \quad (8.2)$$

which leads to the ordinary differential equation

$$\left(1 - \frac{c^2}{c_m^2} \right)^{\beta_0} \frac{dc}{d\psi} = \frac{1}{R} \frac{c}{c_m},$$

whose solution is

$$\psi = Rc_m \int \frac{dc}{c} \left(1 - \frac{c^2}{c_m^2} \right)^{\beta_0} = \frac{1}{2} Rc_m \int \frac{d\tau}{\tau} (1 - \tau)^{\beta_0}. \quad (8.3)$$

To find the velocity field we can use (4.1) which gives

$$\begin{aligned} dz &= \frac{1}{c} e^{i\theta} \left(d\phi + i \frac{P_0}{P} d\psi \right) \\ &= \frac{1}{c} e^{i\theta} \left\{ Rc_m d\theta + i \frac{P_0}{P} c_m R \frac{dc}{c} \frac{P}{P_0} \right\} \\ &= -ic_m R \left\{ \frac{i e^{i\theta} d\theta}{c} - \frac{e^{i\theta} dc}{c^2} \right\}, \end{aligned}$$

so that

$$dz = -ic_m R d \left(\frac{e^{i\theta}}{c} \right)$$

and
$$z = -ic_m R \frac{e^{i\theta}}{c}. \quad (8.4)$$

Since
$$r = |z| = \frac{c_m R}{|c|},$$

it follows that the curves of constant velocity are circles; when $c = c_m$, $|z| = R$ is the limit circle for the motion. Thus using

$$\phi = c_m R \theta, \quad \frac{P}{P_0} \frac{dc}{d\psi} = \frac{c}{Rc_m}, \quad (8.5)$$

we have
$$\frac{\partial}{\partial \phi} = \frac{1}{c_m R} \frac{\partial}{\partial \theta}, \quad \frac{\partial}{\partial \psi} = \frac{c P_0}{R c_m P} \frac{\partial}{\partial c}. \quad (8.6)$$

Substituting in the general differential equation (7.10) we obtain

$$\begin{aligned} & -\frac{1}{a^2 c^2} \frac{\partial^2 \alpha}{\partial t^2} - \frac{2}{a^2 c_m R} \frac{\partial^2 \alpha}{\partial t \partial \theta} - \frac{1}{c_m R} \frac{\partial \alpha}{\partial t} \frac{\partial}{\partial \theta} \left(\frac{1}{a^2} \right) \\ & + \frac{1}{c_m^2 R^2} \frac{\partial}{\partial \theta} \left\{ \left(1 - \frac{c^2}{a^2} \right) \frac{\partial \alpha}{\partial \theta} \right\} + \frac{c P_0}{R c_m P} \frac{\partial}{\partial c} \left\{ \frac{P^2}{P_0^2} \left[\frac{c P_0}{R c_m P} \frac{\partial \alpha}{\partial c} + \beta \right] \right\} = 0. \end{aligned} \quad (8.7)$$

If we assume that the inner boundary is oscillating radially with frequency n the motion at the boundary will be

$$r = r_0 + \delta_0 \cos nt \quad (r_0 > R).$$

Hereafter we shall be concerned principally with the outward propagation of this disturbance. Suppose therefore that we determine a solution for α of the form

$$\alpha = \alpha(c) \frac{\cos}{\sin} \left\{ nt, \right. \quad (8.8)$$

where $\alpha(c)$ is independent of θ . The function $\alpha(c)$ being independent of θ , (8.7) becomes (a also being independent of θ)

$$+ \frac{n^2}{a^2 c^2} \alpha(c) \frac{\cos}{\sin} nt + \frac{c P_0}{R c_m P} \frac{\partial}{\partial c} \left\{ \frac{c P}{R c_m P_0} \frac{d\alpha}{dc} \frac{\cos}{\sin} nt + \frac{P^2 \beta}{P_0^2} \right\} = 0. \quad (8.9)$$

To determine a solution of type (8.8) we must choose β independent of θ in the form

$$\beta = \beta(c) \frac{\cos}{\sin} \left\{ nt, \right.$$

and from (4.8), namely,
$$\frac{\partial \beta}{\partial t} + c^2 \frac{1}{c_m R} \frac{\partial \beta}{\partial \theta} = 0,$$

it follows that β is identically zero, so that (8.9) becomes

$$\frac{n^2 \alpha}{a^2 c^2} + \frac{c P_0}{(R c_m)^2 P} \frac{d}{dc} \left\{ \frac{c P}{P_0} \frac{\partial \alpha}{\partial c} \right\} = 0, \quad (8.10)$$

and the unsteady motion developed in the fluid is then irrotational. If we transform to the variable τ defined in (6.12) this equation may be written

$$\frac{P}{P_0} \frac{(Rc_m n)^2}{a^2 c^2} \alpha + c \frac{d}{dc} \left\{ \frac{P}{P_0} c \frac{\partial \alpha}{\partial c} \right\} = 0,$$

and since

$$c \frac{d}{dc} = 2\tau \frac{d}{d\tau},$$

this becomes

$$\left. \begin{aligned} & \frac{(1-\tau)^{\beta_0} (Rc_m n)^2}{a_0^2 (1-\tau) c_m^2 \tau} \alpha + 4\tau \frac{d}{d\tau} \left\{ (1-\tau)^{\beta_0} \tau \frac{d\alpha}{d\tau} \right\} = 0, \\ & (1-\tau)^{\beta_0} \tau^2 \frac{d^2 \alpha}{d\tau^2} + \frac{d\alpha}{d\tau} \{ \tau(1-\tau)^{\beta_0} - \beta_0 \tau^2 (1-\tau)^{\beta_0-1} \} + \left(\frac{Rn}{2a_0} \right)^2 \frac{(1-\tau)^{\beta_0-1}}{\tau} \alpha = 0, \\ & (1-\tau) \tau^2 \alpha'' + \tau \{ 1 - (1+\beta_0) \tau \} \alpha' + \frac{\lambda^2 \alpha}{\tau} = 0, \end{aligned} \right\} \quad (8.11)$$

where

$$\lambda = \frac{Rn}{2a_0}. \quad (8.12)$$

The unsteady perturbation components for this motion are given by (4.10). Using the transformation (8.6) we derive

$$u_1 = c \frac{\partial \alpha}{\partial \phi} = \frac{c}{c_m R} \frac{\partial \alpha}{\partial \theta} = 0, \quad (8.13)$$

that is, the perturbation component around the circles $|z| = \text{constant}$ is zero. The radial perturbation component is

$$\begin{aligned} v_1 &= \frac{Pc}{P_0} \left(\frac{\partial \alpha}{\partial \psi} + \beta \right) = \frac{Pc}{P_0} \frac{c P_0}{Rc_m P} \frac{\partial \alpha}{\partial c} \\ &= \frac{c^2}{Rc_m} \frac{d\alpha}{dc} \left(= -\frac{d\alpha}{dr} \right), \end{aligned} \quad (8.14)$$

α being defined by (8.8).

9. SOLUTION FOR α WHEN THE BASIC FLOW IS SMALL COMPARED WITH SPEED OF SOUND

If we assume that the basic flow is small compared with the speed of sound (the motion being that at a considerable distance from the oscillating cylinder), then we can assume τ is small in (8.11) and it becomes approximately

$$\tau^2 \alpha'' + \tau \alpha' + \frac{\lambda^2 \alpha}{\tau} = 0. \quad (9.1)$$

Writing

$$\xi = \frac{1}{\tau^{\frac{1}{2}}}, \quad d\xi = -\frac{(\frac{1}{2}) d\tau}{\tau^{\frac{3}{2}}}$$

$$\frac{d}{d\xi} = -2\tau^{\frac{1}{2}} \frac{d}{d\tau}, \quad \text{therefore} \quad \xi \frac{d}{d\xi} = -2\tau \frac{d}{d\tau},$$

$$\tau \frac{d}{d\tau} \left(\tau \frac{d}{d\tau} \right) = \frac{1}{4} \xi \frac{d}{d\xi} \left(\xi \frac{d}{d\xi} \right) = \frac{1}{4} \left\{ \xi^2 \frac{d^2}{d\xi^2} + \xi \frac{d}{d\xi} \right\},$$

$$\tau^2 \frac{d^2}{d\tau^2} + \tau \frac{d}{d\tau} = \frac{1}{4} \left\{ \xi^2 \frac{d^2}{d\xi^2} + \xi \frac{d}{d\xi} \right\},$$

the equation (9.1) becomes

$$\xi^2 \frac{d^2 \alpha}{d\xi^2} + \xi \frac{d\alpha}{d\xi} + 4\lambda^2 \xi^2 \alpha = 0. \quad (9.2)$$

The solutions of (9.2) are $J_0(2\lambda\xi)$ and $Y_0(2\lambda\xi)$. Since small values of τ are assumed, large values of ξ will prevail in (9.2), and we can replace J_0 and Y_0 by their asymptotic expansions, namely,

$$J_0(2\lambda\xi) \sim \left(\frac{1}{\pi\lambda\xi}\right)^{\frac{1}{2}} \cos\{2\lambda\xi - \frac{1}{4}\pi\}, \quad Y_0(2\lambda\xi) \sim \left(\frac{1}{\pi\lambda\xi}\right)^{\frac{1}{2}} \sin\{2\lambda\xi - \frac{1}{4}\pi\}, \quad (9.3)$$

and the complete solution for α will be

$$\alpha = K_1 \tau^{\frac{1}{2}} \cos\left\{\frac{2\lambda}{\tau^{\frac{1}{2}}} - \frac{1}{4}\pi\right\} \frac{\cos}{\sin} nt + K_2 \tau^{\frac{1}{2}} \sin\left\{\frac{2\lambda}{\tau^{\frac{1}{2}}} - \frac{1}{4}\pi\right\} \frac{\cos}{\sin} nt,$$

K_1, K_2 being arbitrary constants. Since

$$\tau = \frac{c^2}{c_m^2} \quad \text{and} \quad \frac{c}{c_m} = \frac{R}{|z|} = \frac{R}{r} = \tau^{\frac{1}{2}}, \quad (9.4)$$

this becomes

$$\alpha = K_1 \sqrt{\frac{R}{r}} \cos\left\{\frac{2\lambda r}{R} - \frac{1}{4}\pi\right\} \frac{\cos}{\sin} nt + K_2 \sqrt{\frac{R}{r}} \sin\left\{\frac{2\lambda r}{R} - \frac{1}{4}\pi\right\} \frac{\cos}{\sin} nt, \quad (9.5)$$

and evidently by suitable combinations of the constants, progressive wave solutions may be formed of the type

$$r^{-\frac{1}{2}} \frac{\cos}{\sin} \left\{ \frac{2\lambda r}{R} - nt \right\}.$$

From (8.12)

$$\frac{2\lambda}{R} = \frac{n}{a_0},$$

so that the progressive wave types are of the form

$$r^{-\frac{1}{2}} \frac{\cos}{\sin} \left\{ n \left(\frac{r}{a_0} - t \right) \right\}. \quad (9.6)$$

Thus, progressive wave types will be propagated with the velocity of sound at a stagnation point, a result to be expected for low basic speeds. It will be observed that the amplitude decreases as $r^{-\frac{1}{2}}$. These waves of course are ordinary sound waves.

10. THE FORM OF THE GENERAL SOLUTION FOR α . CHARACTERISTIC SOLUTIONS

Equation (8.11) may be transformed using

$$\tau = \frac{2\lambda}{z}, \quad (10.1)$$

and it then assumes the form

$$(z^2 - 4\lambda^2) \frac{d^2 \alpha}{dz^2} + \left\{ z + \frac{4\lambda^2(2\beta_0 - 1)}{z} \right\} \frac{d\alpha}{dz} + z^2 \alpha = 0. \quad (10.2)$$

If we write $b = \lambda^2(2\beta_0 - 1) = 4\lambda^2$ for air when $\beta_0 = \frac{5}{2}$, (10.3)

and assume $\alpha = \sum_{r=-\infty}^{\infty} A_{2r} J_{\nu+2r}(z)$, (10.4)

we obtain, assuming that ν does not take the values $\pm 1, \pm 3, \pm 5, \dots$,

$$z^2 \alpha'' + z \alpha' + z^2 \alpha = \sum_{-\infty}^{\infty} (\nu + 2r)^2 A_{2r} J_{\nu+2r}(z),$$

$$\alpha'' = \sum_{-\infty}^{\infty} \frac{1}{4} A_{2r} \{J_{\nu+2r-2} + J_{\nu+2r+2} - 2J_{\nu+2r}\},$$

$$\frac{1}{z} \alpha' = \sum_{-\infty}^{\infty} \frac{1}{4} A_{2r} \left\{ \frac{J_{\nu+2r-2}}{\nu+2r-1} - \frac{2J_{\nu+2r}}{(\nu+2r)^2-1} - \frac{J_{\nu+2r+2}}{\nu+2r+1} \right\}.$$

Thus the following identity has to be satisfied:

$$0 \equiv \sum_{-\infty}^{\infty} (\nu + 2r)^2 A_{2r} J_{\nu+2r} - \lambda^2 \sum_{-\infty}^{\infty} A_{2r} \{J_{\nu+2r-2} + J_{\nu+2r+2} - 2J_{\nu+2r}\} \\ + b \sum_{-\infty}^{\infty} A_{2r} \left\{ \frac{J_{\nu+2r-2}}{\nu+2r-1} - \frac{2J_{\nu+2r}}{(\nu+2r)^2-1} - \frac{J_{\nu+2r+2}}{\nu+2r+1} \right\}.$$

Equating to zero the term in $J_{\nu+2r}$, we obtain a recurrence formula for the A_{2r} coefficients, namely,

$$\left\{ 2\lambda^2 + (\nu + 2r)^2 - \frac{2b}{(\nu + 2r)^2 - 1} \right\} A_{2r} - \lambda^2 \{A_{2r-2} + A_{2r+2}\} + b \left\{ \frac{A_{2r+2}}{\nu + 2r + 1} - \frac{A_{2r-2}}{\nu + 2r - 1} \right\} = 0, \\ \text{or} \quad -\lambda^2 A_{2r-2} \left(1 + \frac{4}{\nu + 2r - 1} \right) + A_{2r} \left[2\lambda^2 + (\nu + 2r)^2 - \frac{8\lambda^2}{(\nu + 2r)^2 - 1} \right] \\ + \lambda^2 A_{2r+2} \left(-1 + \frac{4}{\nu + 2r + 1} \right) = 0, \quad (10.5)$$

where we have now taken $b = 4\lambda^2$. If we take $r = -2, \dots, +3$, the set of relations (10.5) becomes

$$\left. \begin{aligned} -\lambda^2 A_{-6} \left(1 + \frac{4}{\nu - 5} \right) + A_{-4} \left[2\lambda^2 + (\nu - 4)^2 - \frac{8\lambda^2}{(\nu - 4)^2 - 1} \right] + \lambda^2 A_{-2} \left(-1 + \frac{4}{\nu - 3} \right) &= 0, \quad (a) \\ -\lambda^2 A_{-4} \left(1 + \frac{4}{\nu - 3} \right) + A_{-2} \left[2\lambda^2 + (\nu - 2)^2 - \frac{8\lambda^2}{(\nu - 2)^2 - 1} \right] + \lambda^2 A_0 \left(-1 + \frac{4}{\nu - 1} \right) &= 0, \quad (b) \\ -\lambda^2 A_{-2} \left(1 + \frac{4}{\nu - 1} \right) + A_0 \left[2\lambda^2 + (\nu)^2 - \frac{8\lambda^2}{\nu^2 - 1} \right] + \lambda^2 A_2 \left(-1 + \frac{4}{\nu + 1} \right) &= 0, \quad (c) \\ -\lambda^2 A_0 \left(1 + \frac{4}{\nu + 1} \right) + A_2 \left[2\lambda^2 + (\nu + 2)^2 - \frac{8\lambda^2}{(\nu + 2)^2 - 1} \right] + \lambda^2 A_4 \left(-1 + \frac{4}{\nu + 3} \right) &= 0, \quad (d) \\ -\lambda^2 A_2 \left(1 + \frac{4}{\nu + 3} \right) + A_4 \left[2\lambda^2 + (\nu + 4)^2 - \frac{8\lambda^2}{(\nu + 4)^2 - 1} \right] + \lambda^2 A_6 \left(-1 + \frac{4}{\nu + 5} \right) &= 0, \quad (e) \\ -\lambda^2 A_4 \left(1 + \frac{4}{\nu + 5} \right) + A_6 \left[2\lambda^2 + (\nu + 6)^2 - \frac{8\lambda^2}{(\nu + 6)^2 - 1} \right] + \lambda^2 A_8 \left(-1 + \frac{4}{\nu + 7} \right) &= 0, \quad (f) \end{aligned} \right\} \quad (10.6)$$

This infinite set of relations will define an infinite determinant in ν whose vanishing will lead to a set of characteristic solutions for ν . Convergence of the infinite series (10.4) will be assured, since for large values of r equation (10.5) becomes

$$-\lambda^2(A_{2r-2} + A_{2r+2}) + A_{2r}\{2\lambda^2 + (\nu + 2r)^2\} = 0,$$

which is similar to that for the Mathieu coefficients, hence we can accept the convergence as a result of Mathieu function theory.

In order to investigate the infinite determinant of the coefficients A_{2r} we can approach the problem as in Hill's determinant. To ensure convergence of the infinite determinant we divide throughout the r th row by the coefficient

$$2\lambda^2 + (\nu + 2r)^2 - \frac{8\lambda^2}{(\nu + 2r)^2 - 1}$$

of A_{2r} in (10.5), so that (10.5) becomes

$$\frac{-\lambda^2\left(1 + \frac{4}{\nu + 2r - 1}\right)A_{2r-2}}{2\lambda^2 + (\nu + 2r)^2 - \frac{8\lambda^2}{(\nu + 2r)^2 - 1}} + A_{2r} + \frac{\lambda^2\left(-1 + \frac{4}{\nu + 2r + 1}\right)A_{2r+2}}{2\lambda^2 + (\nu + 2r)^2 - \frac{8\lambda^2}{(\nu + 2r)^2 - 1}} = 0. \quad (10.7)$$

Suppose we write
$$\epsilon_{2r} = \frac{-\lambda^2\left(1 + \frac{4}{\nu + 2r - 1}\right)}{2\lambda^2 + (\nu + 2r)^2 - \frac{8\lambda^2}{(\nu + 2r)^2 - 1}}, \quad (10.8)$$

$$\eta_{2r} = \frac{\lambda^2\left(-1 + \frac{4}{\nu + 2r + 1}\right)}{2\lambda^2 + (\nu + 2r)^2 - \frac{8\lambda^2}{(\nu + 2r)^2 - 1}}, \quad (10.9)$$

then
$$\epsilon_{2r}A_{2r-2} + A_{2r} + \eta_{2r}A_{2r+2} = 0, \quad (10.10)$$

and the determinant of the coefficients is then

$$\Delta(\nu) = \begin{vmatrix} \epsilon_{-6} & 1 & \eta_{-6} & 0 & 0 & 0 & 0 & 0 \\ 0 & \epsilon_{-4} & 1 & \eta_{-4} & 0 & 0 & 0 & 0 \\ 0 & 0 & \epsilon_{-2} & 1 & \eta_{-2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \epsilon_0 & 1 & \eta_0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \epsilon_2 & 1 & \eta_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \epsilon_4 & 1 & \eta_4 \\ & & 0 & 0 & 0 & 0 & \epsilon_6 & 1 \end{vmatrix}. \quad (10.11)$$

Write $-\nu$ for ν , then

$$\epsilon_{-2r} = \frac{+\lambda^2\left(-1 + \frac{4}{\nu + 2r + 1}\right)}{2\lambda^2 + (\nu + 2r)^2 - \frac{8\lambda^2}{[(\nu + 2r)^2 - 1]}} = \eta_{2r},$$

and

$$\eta_{-2r} = \frac{\lambda^2 \left(-1 + \frac{4}{-\nu - 2r + 1} \right)}{2\lambda^2 + (\nu + 2r)^2 - \frac{8\lambda^2}{(\nu + 2r)^2 - 1}} = \epsilon_{2r},$$

$$\Delta(-\nu) = \begin{vmatrix} \eta_6 & 1 & \epsilon_6 & 0 & 0 & 0 & 0 & 0 \\ 0 & \eta_4 & 1 & \epsilon_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & \eta_2 & 1 & \epsilon_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \eta_0 & 1 & \epsilon_0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \eta_{-2} & 1 & \epsilon_{-2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \eta_{-4} & 1 & \epsilon_{-4} \end{vmatrix},$$

and since the determinant is infinite we have $\Delta(\nu) = \Delta(-\nu)$, that is, $\Delta(\nu)$ is an even function of ν . Replacing r by $r-1$ and ν by $\nu+2$,

$$\epsilon_{2r-2}(\nu+2) = \epsilon_{2r}(\nu), \quad \eta_{2r-2}(\nu+2) = \eta_{2r}(\nu),$$

and thus $\Delta(\nu+2) = \Delta(\nu)$, and similarly $\Delta(\nu \pm 2m) = \Delta(\nu)$, hence $\Delta(\nu)$ has period $\nu = 2$.

The singularities of $\Delta(\nu)$ occur as simple poles of the terms ϵ_{2r} , η_{2r} . We note that $\nu = \pm(2r-1)$ are not simple poles of the functions, and from (10.8) and (10.9) we derive the simple poles at the zeroes of the function

$$2\lambda^2 + (\nu + 2r)^2 - \frac{8\lambda^2}{(\nu + 2r)^2 - 1} = 0,$$

which leads to the two expressions

$$\left. \begin{aligned} (\nu + 2r)^2 &= \frac{1}{2} \{ \sqrt{(4\lambda^4 + 36\lambda^2 + 1)} - (2\lambda^2 - 1) \} = k_1^2(\lambda), \\ (\nu + 2r)^2 &= -\frac{1}{2} \{ \sqrt{(4\lambda^4 + 36\lambda^2 + 1)} + 2\lambda^2 - 1 \} = -k_2^2(\lambda), \end{aligned} \right\} \quad (10.12)$$

where $k_1^2(\lambda)$ and $k_2^2(\lambda)$ are positive quantities. Hence the simple poles are

$$\left. \begin{aligned} \nu &= -2r \pm k_1(\lambda) & (r = -\infty, \infty), \\ \nu &= -2r \pm ik_2(\lambda) & (r = -\infty, \infty), \end{aligned} \right\} \quad (10.13)$$

that is, one set along the real axis, a second set along the lines $\pm ik_2(\lambda)$.

It is easily shown that $\Delta(\nu)$ is an analytic function except for the poles (10.13), and these poles are simple provided $\lambda > 0$.

The function $\chi(\nu)$ defined by

$$\chi(\nu) = A \{ \cot \frac{1}{2} \pi (\nu - k_1) - \cot \frac{1}{2} \pi (\nu + k_1) \} + B \{ \cot \frac{1}{2} \pi (\nu - ik_2) - \cot \frac{1}{2} \pi (\nu + ik_2) \} \quad (10.14)$$

is also an analytic function except for simple poles at the points (10.13), and it has period $\nu = 2$. It follows that the function

$$f(\nu) = \Delta(\nu) - \chi(\nu) \quad (10.15)$$

will be of period 2. Furthermore, if A and B are so chosen that the singularities at $\nu = k_1$ and $\nu = ik_2$ cancel, then in virtue of the periodicity it follows that all the singularities of $f(\nu)$ disappear and $f(\nu)$ will be an analytic function. The singularities at $\nu = k_1$ and $\nu = ik_2$ will cancel provided that the residues at these poles are the

same, and it is evident that this method leads to two compatible linear equations for A and B whose existence is thereby proved. Since $\Delta(\nu)$ and $\chi(\nu)$ are bounded as ν tends to infinity, it follows that $f(\nu)$ is bounded, and thus by Liouville's Theorem $f(\nu)$ is a constant. To determine this constant make $\nu \rightarrow i\infty$ in (10.15), then since $\lim_{\nu \rightarrow i\infty} \Delta(\nu) = 1$ and $\lim_{\nu \rightarrow i\infty} \chi(\nu) = 0$ it follows that $f(\nu) \equiv 1$. Thus

$$\Delta(\nu) - 1 = \chi(\nu). \quad (10.16)$$

The method outlined above for the determination of A and B may be avoided if we note that (10.16) is an identity in ν and A and B may be determined alternatively by equating the left- and right-hand sides at any two convenient points in the ν -plane. These points are taken to be $\nu = 0$ and $\nu = 1$. With $\nu = 0$ we obtain the equation

$$\Delta(0) - 1 = -2A \cot \frac{1}{2}\pi k_1 + 2iB \coth \frac{1}{2}\pi k_2, \quad (10.17)$$

and when $\nu = 1$ the equation

$$\Delta(1) - 1 = 2A \tan \frac{1}{2}\pi k_1 + 2iB \tanh \frac{1}{2}\pi k_2. \quad (10.18)$$

When $\nu = 1$ it follows from (10.8) and (10.9) that

$$\epsilon_0 = \eta_{-2} = 1, \quad \eta_0 = \eta_2 = \epsilon_{-2} = 0,$$

and thus $\Delta(1) = 0$. The corresponding values $\nu = \pm 1, \pm 3, \pm 5, \dots$ which make $\Delta(\nu)$ vanish are excluded as solutions of the problem (see 10.4 et seq.). Equation (10.18) now becomes

$$-1 = 2A \tan \frac{1}{2}\pi k_1 + 2iB \tanh \frac{1}{2}\pi k_2, \quad (10.19)$$

and (10.17), (10.19) lead to

$$\left. \begin{aligned} A\{\cosh \pi k_2 - \cos \pi k_1\} &= -\frac{1}{2} \sin(\pi k_1) \{\Delta(0) \sinh^2(\frac{1}{2}\pi k_2) + 1\}, \\ \text{and } iB\{\cosh \pi k_2 - \cos \pi k_1\} &= \frac{1}{2} \sinh(\pi k_2) \{\Delta(0) \sin^2(\frac{1}{2}\pi k_1) - 1\}. \end{aligned} \right\} \quad (10.20)$$

The equation $\Delta(\nu) = 0$, which gives the characteristic values, becomes $\chi(\nu) = -1$, and this reduces to

$$\frac{2A \sin \pi k_1}{\cos \pi k_1 - \cos \nu \pi} + \frac{2iB \sinh \pi k_2}{\cosh \pi k_2 - \cos \nu \pi} = -1. \quad (10.21)$$

The only difficulty in the solution of this transcendental equation lies in the evaluation of $\Delta(0)$ for any specified λ .

In the case of sufficiently small values of λ (that is, when the frequency of the oscillating cylinder is sufficiently small) a solution for ν of the equation (10.21) may be obtained as follows. The following formulae have been derived and are correct to the final power included:

$$\left. \begin{aligned} \Delta(0) &= 1 + \frac{3}{4}\lambda^2 + \frac{43}{288}\lambda^4 + \dots, \\ k_1 &= 1 + 4\lambda^2 - 48\lambda^4 + 912\lambda^6 + \dots, \\ k_2 &= \lambda \sqrt{(10) \{1 - 4\lambda^2 + 64\lambda^4 + \dots\}}, \\ \cos \pi k_1 &= -1 + 8\pi^2\lambda^4 - 192\pi^2\lambda^6 + \dots, \\ \sin \pi k_1 &= -4\pi\lambda^2 + 48\pi\lambda^4 - \lambda^6(912\pi + \frac{32}{3}\pi^3) + \dots, \\ \cosh \pi k_2 &= 1 + 5\pi^2\lambda^2 + \lambda^4(\frac{25}{6}\pi^4 - 40\pi^2) + \dots, \\ \sinh \pi k_2 &= \pi\lambda \sqrt{(10) \{1 + \lambda^2(\frac{5}{3}\pi^2 - 4) + \lambda^4(64 - 20\pi^2) + \dots\}}, \\ A &= \pi\lambda^2 \{1 - 12\lambda^2 + \dots\}, \\ iB &= \frac{3}{16}\pi\lambda^3 \sqrt{(10) \{1 - \lambda^2(\frac{821}{216} + \frac{37}{6}\pi^2) \dots\}}, \end{aligned} \right\} \quad (10.22)$$

and the solution for ν is obtained from

$$-1 = \frac{-8\pi^2\lambda^4(1-12\lambda^2)^2}{-1+8\pi^2\lambda^4-\cos\nu\pi} + \frac{\frac{15}{4}\pi^2\lambda^4[1-\lambda^2(\frac{1685}{216}+\frac{9}{2}\pi^2)\dots]}{1+5\pi^2\lambda^2+\lambda^4(\frac{25}{6}\pi^4-40\pi^2)-\cos\nu\pi}. \quad (10.23)$$

It will be noted that $\nu = 1$ satisfies this equation to order λ^4 , and it is easily shown that the second solution is

$$\nu = i\lambda\sqrt{(10)}\{1-\frac{29}{8}\lambda^2+\dots\}, \quad (10.24)$$

so that for small λ the solution for ν is an imaginary quantity. There will be an infinite set of solutions for ν which make $\Delta(\nu)$ vanish, and these will be obtained from (10.24) by the addition or subtraction of an even integer.

In order to check the result (10.23) we return to the equations (10.6) and solve for ν by an alternative method (this method is possible, as in Mathieu function theory, for sufficiently small values of λ). When A_{-4} is neglected in (10.6(b)), this equation defines A_{-2} in terms of A_0 . Similarly, by neglecting A_4 in (10.6(d)), this equation defines A_2 in terms of A_0 . Substituting for A_2 and A_{-2} in (10.6(c)) and equating to zero, the coefficient of A_0 then gives an approximate form for the characteristic equation $\Delta(\nu) = 0$. This equation is

$$\frac{-\lambda^4\left(1-\frac{16}{(\nu-1)^2}\right)}{2\lambda^2+(\nu-2)^2-\frac{8\lambda^2}{(\nu-2)^2-1}} + 2\lambda^2+\nu^2-\frac{8\lambda^2}{\nu^2-1} - \frac{\lambda^4\left(1-\frac{16}{(\nu+1)^2}\right)}{2\lambda^2+(\nu+2)^2-\frac{8\lambda^2}{(\nu+2)^2-1}} = 0, \quad (10.25)$$

and for small λ it is easily verified that (10.24) is a solution of this equation. This alternative method leads to the solution for ν more rapidly than the original method, but suffers from the disadvantage that it may be applied for small values of λ only.

With the solution (10.24) for ν we obtain

$$A_{-2} = \frac{1}{4}\lambda^2 A_0(5+9i\lambda\sqrt{(10)}\dots), \quad A_2 = \frac{1}{4}\lambda^2 A_0(5-9i\lambda\sqrt{(10)}\dots), \quad (10.26)$$

so that A_{-2} and A_2 are, to this order of magnitude in λ , conjugate complex quantities. Hereafter the terms involving A_{2r} will be neglected in the general solution (10.4), and in this case, for sufficiently small values of λ , the solution for α is

$$\alpha = A_0 J_{i\lambda\sqrt{(10)}}(z) e^{int}. \quad (10.27)$$

Two important cases may be considered:

$$(i) \text{ Since } z = \frac{2\lambda}{\tau}, \quad (10.28)$$

when λ is small, z will be small provided τ is near its upper limit, namely, $\tau = 1$. In this case we may approximate to $J_\nu(z)$ by its leading term, and we thus obtain the solution

$$\begin{aligned} \alpha &= e^{int} \left(\frac{1}{2}z\right)^{i\sqrt{(10)}\lambda} \\ &= \exp\{i[nt + \lambda\sqrt{(10)} \log(\frac{1}{2}z)]\}. \end{aligned} \quad (10.29)$$

This solution corresponds to a system of converging waves, and the more appropriate diverging wave solution arises from the second solution $J_{-\nu}(z)$, which leads to

$$\alpha = \exp\{i[nt - \lambda\sqrt{(10)} \log(\frac{1}{2}z)]\}. \quad (10.30)$$

It easily follows from this that the disturbance is propagated at the distance r with the velocity V_r given by

$$V_r = \frac{nr}{\lambda\sqrt{(10)}} = \frac{a_0}{1.557} \frac{r}{R} \quad (\text{therefore } V_r < a_0), \quad (10.31)$$

this result being valid for values of r near R and for small values of λ .

(ii) The second case of (10.27) which can be easily treated is the case of small λ and large z which arises when τ is near its lower limit, namely, $\tau = 0$. The conditions approach the still-air propagation problem, and the velocity of propagation should be a_0 . Using the asymptotic solution for $J_\nu(z)$ for large z we derive

$$J_{i\lambda\sqrt{(10)}}(z) \sim \left(\frac{2}{\pi z}\right)^{\frac{1}{2}} \cos\left\{z - \frac{1}{2}i\pi\sqrt{(10)}\lambda - \frac{1}{4}\pi\right\}.$$

Since $J_{-i\lambda\sqrt{(10)}}(z)$ is also a solution it follows that

$$\left(\frac{2}{\pi z}\right)^{\frac{1}{2}} e^{\pm i(z - \frac{1}{2}\pi)}$$

are solutions, and the diverging wave solution is given by

$$\left(\frac{2}{\pi z}\right)^{\frac{1}{2}} e^{i(nt-z)} \propto \frac{1}{r^{\frac{1}{2}}} \exp\left[i\left(nt - \frac{nr}{a_0}\right)\right], \quad (10.32)$$

where we have used $z = nr/a_0$. Thus the velocity of propagation of the waves is a_0 as expected.

When λ is large it does not seem likely that an analytic solution to the problem can be derived.

REFERENCES

- Carrier, G. F. & Carlson, F. D. 1946 *Quart. Appl. Math.* **4**, 1-12.
 Lewis, T. 1938 *Quart. J. Math.*, **9**, no. 33, p. 21.

The electron pairing theory of the structure of conjugated hydrocarbons

BY W. MOFFITT, *British Rubber Producers' Research Association,
Welwyn Garden City*

(Communicated by E. K. Rideal, F.R.S.—Received 15 June 1949)

The electron pairing theory of molecular structure has been extended in its application to conjugated hydrocarbons. Formulae for calculating internuclear distances, force constants and interaction constants are derived. These appear to give rather better results than the already successful molecular orbital theory. The interaction of two conjugated systems is discussed, and it is shown how the conjugating power of a hydrocarbon residue may be correlated with its residual affinity

INTRODUCTION

So far, the pairing (or HLSP = Heitler, London, Slater, Pauling) approximation has been successfully applied to the calculation of the resonance energies (Pauling 1933) of the bond lengths and bond energies (Penney 1937) and of the chemical reactivities (Moffitt 1949) of conjugated hydrocarbons. In all these treatments it was tacitly assumed that the variation of the exchange integrals with internuclear distances may be treated as a second-order effect. The only development (Penney & Lloyd 1939) of the theory which permits the calculation of force constants has the disadvantage of considerable computational complexities in as much as the energy is minimized with respect to the bond lengths at the outset.

In the belief that this is in some respects an unnecessary refinement, it will be shown how force constants and interaction constants may be calculated, without difficulty, once the first-order wave functions have been determined on the assumption of equal exchange integrals for neighbouring atoms. The author's aim is the same as that of Coulson & Longuet-Higgins (1948) in their recent simplification of the molecular orbital method as applied to similar systems by Lennard-Jones (1937). In particular, it will be found convenient to introduce certain quantities, called mutual bond polarizabilities, which are closely related to similar quantities, first introduced for the molecular orbital theory by Coulson & Longuet-Higgins.

THE PAIRING METHOD

As is customary in HLSP theory, we facilitate the resolution of the exchange degeneracy by expressing the wave function ψ , representing the unsaturation or π electrons, as a linear combination of simpler structure-functions ψ_α , ν in number, writing

$$\psi = \sum_{\alpha} a_{\alpha} \psi_{\alpha} \quad (\alpha = 1, \dots, \nu). \quad (1)$$

For simplicity, we shall confine ourselves to a homopolar representation in which each of the ψ_α represents a purely covalent structure, though this restriction is not necessary. Thus each ψ_α is an anti-symmetrical many-electron function which may

be uniquely specified by a directed valence-bond pattern; the different ψ_α are linearly independent, normalized, but not in general mutually orthogonal. (In fact, more than ν structure-functions ψ_α may in general be enumerated, but the manifold which they subtend together is of ν dimensions only; the number of possible structures exceeds the number of independent structures. We therefore choose a convenient set of ν independent ψ_α in terms of which to expand ψ . It is easily shown that the observables which we define in the sequel are independent of this choice.)

Since the ψ_α are real, we introduce the notation

$$H_{\alpha\beta} = \int \psi_\alpha \mathcal{H} \psi_\beta dv = H_{\beta\alpha}, \quad S_{\alpha\beta} = \int \psi_\alpha \psi_\beta dv = S_{\beta\alpha}, \quad (2)$$

so that the energy E may be expressed in the form

$$E = \left(\sum_\alpha \sum_\beta a_\alpha H_{\alpha\beta} a_\beta \right) / \left(\sum_\alpha \sum_\beta a_\alpha S_{\alpha\beta} a_\beta \right). \quad (3)$$

The a_α are determined by minimizing this energy with respect to these a_α , which may therefore be treated as variational parameters. The $S_{\alpha\beta}$ are readily determined as non-negative powers of $\frac{1}{2}$ —most conveniently, perhaps, by means of certain rules devised by Pauling (1933)—and the $H_{\alpha\beta}$ may be similarly expressed in the form

$$H_{\alpha\beta} = S_{\alpha\beta} Q + \sum_r S_{\alpha\beta} \gamma_{\alpha\beta}^r J_r \quad (r = 1, \dots, n). \quad (4)$$

Q is the configurational Coulomb energy term, and the J_r are the exchange integrals

$$J_r = \iint \phi_i(1) \phi_j(2) \mathcal{H} \phi_j(1) \phi_i(2) dv_1 dv_2,$$

which are directly related to the strength of the bonding between atoms $i = i(r)$ and $j = j(r)$. It is assumed that the only non-vanishing exchange integrals are those for which i and j are nearest neighbours, and that these are numerically equal to J say. The summation in (4) is therefore taken over all effective bonds $r (= 1, \dots, n)$, or over all pairs of nearest neighbours. If the a_α are chosen so as to normalize ψ , we have

$$E = Q + \sum_r c_r(a) J_r, \quad (5)$$

where the c_r are quadratic forms in the a_α ,

$$c_r = \sum_\alpha \sum_\beta a_\alpha S_{\alpha\beta} \gamma_{\alpha\beta}^r a_\beta. \quad (6)$$

On minimizing the energy, by setting

$$\frac{\partial E}{\partial a_\alpha} = 0 \quad (\alpha = 1, \dots, \nu), \quad (7)$$

we obtain the set of linear equations

$$\sum_\beta (H_{\alpha\beta} - E S_{\alpha\beta}) a_\beta = 0 \quad (\alpha = 1, \dots, \nu), \quad (8)$$

from which we may eliminate the a_β to determine the critical values of the expression on the right-hand side of (3) by means of

$$H_{\alpha\beta} - E S_{\alpha\beta} = 0. \quad (9)$$

Any root E^τ of this equation determines a corresponding normalized wave function ψ^τ , or set of a_α^τ ($\alpha = 1, \dots, \nu$), to the extent of a trivial factor of modulus unity. Since we shall, at first, be only concerned with one particular state of a molecule, we shall drop the superscript τ which distinguishes the ν different 'states' arising from our homopolar representation. In fact, of course, this representation is incomplete, and only certain of the ψ^τ , in particular that ψ^τ which represents the state of lowest energy, may be taken as good approximations to the wave functions of observed states.

We may compare (5) with Dirac's formal solution of the perturbation problem

$$E = Q - \frac{1}{2} \sum_r (1 + 4\overline{\mathbf{s}_i \cdot \mathbf{s}_j}) J_r;$$

$\overline{\mathbf{s}_i \cdot \mathbf{s}_j}$ represents the mutual inclination of the spin vectors of the electrons associated with atoms $i = i(r)$ and $j = j(r)$ bound by the effective bond r . Penney (1937) has defined bond orders for HLSP theory by means of the relation

$$p_r = -\frac{4}{3} \overline{\mathbf{s}_i \cdot \mathbf{s}_j},$$

so that

$$c_r(a) = -\frac{1}{2} + \frac{3}{2} p_r(a), \quad (10)$$

and

$$E = Q - \frac{1}{2} \sum_r J_r + \frac{3}{2} \sum_r p_r(a) J_r. \quad (11)$$

Since the p_r take the values 0 and 1 for pure single and double bonds respectively, we may call

$$\mathcal{E} = \frac{3}{2} \sum_r p_r(a) J_r, \quad (12)$$

the total exchange energy of π -electronic binding—or, simply, the mobile exchange energy. The term $-\left(\frac{1}{2}\right) \sum_r J_r$ is not contained in $\mathcal{E}$, because it would appear even if all the r were formally pure single bonds; in the role of a constant exchange repulsion, it therefore contributes, together with Q , to what may be called the total configurational energy—which is independent of the values assumed by the eigenvalues of the scalar product operators $\mathbf{s}_i \cdot \mathbf{s}_j$. By use of (9), we notice that the total derivative, $D\mathcal{E}/DJ_r$, of the mobile exchange energy with respect to J_r satisfies the relation

$$\frac{D\mathcal{E}}{DJ_r} = \frac{\partial \mathcal{E}}{\partial J_r} + \sum_\alpha \frac{\partial \mathcal{E}}{\partial a_\alpha} \frac{\partial a_\alpha}{\partial J_r} = \frac{\partial \mathcal{E}}{\partial J_r} = \frac{3}{2} p_r. \quad (13)$$

THE MUTUAL BOND POLARIZABILITIES

In order to consider the effect of variations in the values of the different exchange integrals J_r , we shall find it convenient to introduce mutual bond polarizabilities by means of the equation

$$\pi_{r,s} = \frac{\partial p_r}{\partial J_s} \quad (r, s = 1, \dots, n). \quad (14)$$

These will show how the order of some bond r is affected by altering the value of one of the J_s . We note that, since $\partial/\partial J_r$ and $\partial/\partial J_s$ commute,

$$\pi_{r,s} = \frac{\partial p_r}{\partial J_s} = \frac{2}{3} \frac{\partial^2 \mathcal{E}}{\partial J_s \partial J_r} = \frac{\partial p_s}{\partial J_r} = \pi_{s,r}. \quad (15)$$

Further, by differentiating (11) we obtain the identity

$$\sum_s \pi_{r,s} = 0 = \sum_r \pi_{r,s}. \quad (16)$$

A rather important theorem is that, for ground states,

$$\pi_{r,r} < 0; \quad (17)$$

that is, the self-polarizability of a bond is always negative. Since the addition of a positive increment to J_r corresponds to weaker bonding at r , this implies a parallel decrease in p_r ; in this form, (17) is an almost necessary prerequisite for any definition of self-bond-polarizability. It may be proved as follows: We may expand the mobile exchange energy, corresponding to a variation δJ_r , positive or negative, in the value of J_r , as a Taylor's series

$$\mathcal{E} + \frac{\partial \mathcal{E}}{\partial J_r} \delta J_r + \frac{1}{2!} \frac{\partial^2 \mathcal{E}}{\partial J_r^2} (\delta J_r)^2 + \dots = \frac{3}{2} \sum_s p_s J_s + \frac{3}{2} p_r \delta J_r + \frac{3}{4} \pi_{r,r} (\delta J_r)^2 + \dots,$$

and neglect all terms beyond the second power in δJ_r . The first two terms give the energy of the perturbed state, using the same wave function for this state as was used for the unperturbed ground state. The last term, however, allows the electronic motions, or their wave functions, to adapt themselves to the perturbation and, by the variational theorem, is accordingly negative. Since $(\delta J_r)^2$ is essentially positive, δJ_r being real, the theorem follows.

The $\pi_{r,s}$ may be calculated for any given state either by means of perturbation theory or by adopting the following technique which is more convenient when the wave functions have only been determined accurately for complicated molecules in the states of interest (generally their ground states). We assume, then, that the normalized wave function ψ has been determined. Substituting the right-hand side of (5) for E in (8), we differentiate partially with respect to J_r (E may here be regarded as a Lagrangian multiplier ensuring normalization in the variational problem) and obtain

$$\sum_\beta (H_{\alpha\beta} - E S_{\alpha\beta}) \frac{\partial a_\beta}{\partial J_r} + \sum_\beta \frac{\partial H_{\alpha\beta}}{\partial J_r} a_\beta - \left(c_r + \sum_s \frac{\partial c_s}{\partial J_r} J_s \right) \sum_\beta S_{\alpha\beta} a_\beta = 0. \quad (18)$$

But from (6) we have
$$\frac{\partial c_s}{\partial J_r} = 2 \sum_\beta \frac{\partial a_\beta}{\partial J_r} \sum_\lambda S_{\beta\lambda} \gamma_{\beta\lambda}^s a_\lambda, \quad (19)$$

and from (4) and (8)

$$\sum_s \sum_\lambda S_{\beta\lambda} \gamma_{\beta\lambda}^s a_\lambda J_s = (E - Q) \sum_\lambda S_{\beta\lambda} a_\lambda = (E - Q) \sum_\lambda a_\lambda S_{\lambda\beta},$$

so that
$$\sum_s \frac{\partial c_s}{\partial J_r} J_s = 2 \sum_\beta \frac{\partial a_\beta}{\partial J_r} \sum_s \sum_\lambda S_{\beta\lambda} \gamma_{\beta\lambda}^s a_\lambda J_s = 2(E - Q) \sum_\beta \frac{\partial a_\beta}{\partial J_r} \sum_\lambda a_\lambda S_{\lambda\beta}.$$

Accordingly (18) may be rewritten in the form

$$\sum_\beta \mu_{\alpha\beta} \frac{\partial a_\beta}{\partial J_r} + \lambda_\alpha^r = 0 \quad (\alpha = 1, \dots, \nu; r = 1, \dots, n), \quad (20)$$

where

$$\mu_{\alpha\beta} = (H_{\alpha\beta} - E S_{\alpha\beta}) - 2(E - Q) \left(\sum_\kappa S_{\alpha\kappa} a_\kappa \right) \left(\sum_\lambda a_\lambda S_{\lambda\beta} \right), \quad (21)$$

$$\lambda_\alpha^r = \sum_\kappa (\gamma_{\alpha\kappa}^r - c_r) S_{\alpha\kappa} a_\kappa. \quad (22)$$

The linear set of $n\nu$ equations (20) serves to determine the $n\nu$ unknowns $\partial\alpha_\beta/\partial J_r$, from which we may evaluate the $\partial c_s/\partial J_r$ using (19), and the $\pi_{r,s}$ from the relation

$$\pi_{r,s} = \frac{2}{3} \frac{\partial c_s}{\partial J_r}. \quad (23)$$

We notice that the $\mu_{\alpha\beta}$ contain J as a factor, whereas the λ_α^r are dimensionless; the $\pi_{r,s}$ will therefore be given as multiples of the reciprocal of the exchange integral J .

AN EXPLICIT FORM FOR THE POLARIZABILITIES

If the wave functions of all the states ψ^τ arising from our homopolar representation have been determined, the mutual bond polarizabilities are most readily calculated by means of perturbation theory. This method enables us to give explicit formulae for the polarizabilities in terms of the coefficients α_α^r and the E^τ . As mentioned before, superscripts refer to the different, mutually orthogonal, ψ^τ which may be obtained by considering all ν roots of the secular equation (9). Thus $\pi_{r,s}^\tau$ indicates the mutual polarizability of bonds r and s for the state represented by ψ^τ with energy E^τ . It will be assumed that the enumeration of the ψ_α in the expansions

$$\psi^\tau = \sum_\alpha \alpha_\alpha^r \psi_\alpha \quad (\alpha, \tau = 1, \dots, \nu) \quad (24)$$

is the same for all τ .

We consider the effect of altering the value of J_r by a small amount δJ_r . By means of perturbation theory, the wave functions for the perturbed states are given by

$$\psi^\tau + \delta\psi^\tau = \psi^\tau + \sum_\rho \frac{\mathcal{H}'_{r\tau\rho}}{E^\tau - E^\rho} \psi^\rho \quad (\rho = 1, \dots, \tau-1, \tau+1, \dots, \nu), \quad (25)$$

where the values of the perturbation matrix elements are

$$\mathcal{H}'_{r\tau\rho} = \int \psi^\tau \mathcal{H}'_r \psi^\rho dv = \sum_\alpha \sum_\beta \alpha_\alpha^r \alpha_\beta^\rho \int \psi_\alpha \mathcal{H}'_r \psi_\beta dv \quad (26)$$

$$\text{with} \quad \int \psi_\alpha \mathcal{H}'_r \psi_\beta dv = S_{\alpha\beta} \gamma_{\alpha\beta}^r \delta J_r.$$

We introduce the notation

$$c_r^{\tau\rho} = \sum_\alpha \sum_\beta \alpha_\alpha^r S_{\alpha\beta} \gamma_{\alpha\beta}^r \alpha_\beta^\rho, \quad \bar{\alpha}_\alpha^r = \sum_\rho \alpha_\alpha^\rho \frac{c_r^{\tau\rho}}{E^\tau - E^\rho}, \quad (27)$$

so that, on combining equations (25), (26) and (27), we have

$$\delta\psi^\tau = \delta J_r \sum_\rho \frac{c_r^{\tau\rho}}{E^\tau - E^\rho} \psi^\rho = \delta J_r \sum_\alpha \bar{\alpha}_\alpha^r \psi_\alpha. \quad (28)$$

We have therefore determined the first-order wave functions corresponding to our perturbation δJ_r .

Let us next determine the variations with J_r of the values of the c_s , as defined in equations (6), for the state ψ^τ . When δJ_r is small,

$$\begin{aligned} \delta c_s^\tau &= \sum_\alpha \sum_\beta (\alpha_\alpha^r + \delta J_r \bar{\alpha}_\alpha^r) S_{\alpha\beta} \gamma_{\alpha\beta}^s (\alpha_\beta^r + \delta J_r \bar{\alpha}_\beta^r) - \sum_\alpha \sum_\beta \alpha_\alpha^r S_{\alpha\beta} \gamma_{\alpha\beta}^s \alpha_\beta^r \\ &= 2\delta J_r \sum_\alpha \sum_\beta \bar{\alpha}_\alpha^r S_{\alpha\beta} \gamma_{\alpha\beta}^s \alpha_\beta^r. \end{aligned}$$

Substituting from (27), this becomes

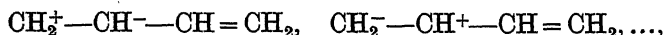
$$\frac{\partial c_s^\tau}{\partial J_r} = 2 \sum_\alpha \sum_\beta \sum'_\rho \frac{c_r^{\tau\rho}}{E^\tau - E^\rho} a_\alpha^\rho S_{\alpha\beta} \gamma_{\alpha\beta}^s a_\beta^\tau = 2 \sum'_\rho \frac{c_r^{\tau\rho} c_s^{\rho\tau}}{E^\tau - E^\rho},$$

whence, from (23),
$$\pi_{r,s}^\tau = \frac{4}{3} \sum'_\rho \frac{c_r^{\tau\rho} c_s^{\rho\tau}}{E^\tau - E^\rho}. \quad (29)$$

This result has certain interesting implications. In the first place, if ψ^τ refers to the ground state, $E^\tau < E^\rho$ and

$$\pi_{r,r}^\tau = \frac{4}{3} \sum'_\rho \frac{(c_r^{\tau\rho})^2}{E^\tau - E^\rho} < 0,$$

furnishing us with a more direct proof of our theorem (17). It also brings to light a limitation of the method which is inherent in the use of a purely covalent representation of the different states. In general it may be said that, whereas such a homopolar representation is found to give good results for the ground states of unsaturated hydrocarbons, the descriptions of the excited states which it offers are not reliable unless supplementary ionic structures are included. Thus the wave functions of the excited states B_s , B_u of *cis*- and *trans*-butadiene respectively are easily shown to contain no purely covalent component structures. Since this limits the 'flexibility' of our conjugated systems unduly, particularly when these are small, we may expect that the perturbed wave functions (25) lead to incorrect values for the $\pi_{r,s}^\tau$, even when ψ^τ refers to the ground state. For example, in butadiene it is found that, if the bonds are indexed 1, 2, 3 in that order along the chain, then the simple theory which we have pursued gives $\pi_{1,1} = \pi_{1,3}$; but a more detailed analysis, including the contribution of structures



would show that $|\pi_{1,1}| > |\pi_{1,3}|$. However, the number of independent covalent structures increases rapidly with the number of carbon atoms, so that we feel reasonably confident in our predictions for the ground states of molecules containing six or more conjugating atoms. This confidence is later justified by our calculations on the benzene molecule.

Lastly, let us consider the effect of a molecule with external symmetry. It is well known that sets of the ψ_α may be linearly combined to form irreducible representations of the molecular point group, by means of which the secular equation (9) is readily factorized and its roots more simply extracted. Similarly if, on changing the value of J_r (by lengthening bond r , say), the molecule still retains certain non-trivial symmetry elements, the calculation of the $\pi_{r,s}^\tau$ may be facilitated. For the ψ^ρ may always be chosen so as to transform like irreducible representations of the group of lower symmetry—and it is then easily shown that the $c_r^{\tau\rho}$ will vanish unless ψ^ρ and ψ^τ transform in the same way. This simplifies the calculation of the mutual bond polarizabilities for the ground state of the benzene molecule, for example, and would be particularly important if ionic structures were to be included in the analysis.

THE VIBRATIONAL POTENTIAL ENERGY

Let us now return to the general theory of variations in the individual bond lengths and consider the dynamics of molecular vibrations. As usual, we begin by dividing the appropriate Hamiltonian, obtained by separating the nuclear from the electronic motions, into a term $\mathcal{T}$, representing the kinetic energy of the nuclei, and a term $\mathcal{V}$. The potential energy term $\mathcal{V}$ is uniquely determined by the nuclear configuration. We shall specify the nuclear co-ordinates by means of the bond lengths x_r ($r = 1, \dots, n$), together with as many other co-ordinates x_t ($t = n+1, \dots, m$) as are necessary to complete this specification. $\mathcal{V}$ may then be expanded as a Taylor's series about any, say the equilibrium, configuration $\{\bar{x}_u\}$, ($u = 1, \dots, m$), as follows:

$$\mathcal{V}(x) - \mathcal{V}(\bar{x}) = \sum_u \frac{\partial \mathcal{V}}{\partial \bar{x}_u} (x_u - \bar{x}_u) + \frac{1}{2!} \sum_u \sum_v \frac{\partial^2 \mathcal{V}}{\partial \bar{x}_u \partial \bar{x}_v} (x_u - \bar{x}_u) (x_v - \bar{x}_v) + \dots, \quad (30)$$

where $\partial \mathcal{V} / \partial \bar{x}_u$ represents the value of $\partial \mathcal{V} / \partial x_u$ in the given configuration $\{\bar{x}_u\}$. When, as we shall henceforth suppose, $\{\bar{x}_u\}$ does indeed represent the equilibrium configuration, the first terms on the right-hand side of (30) vanish, so that for small displacements from equilibrium, we have

$$\mathcal{V}(x) - \mathcal{V}(\bar{x}) = \frac{1}{2} \sum_u \sum_v \frac{\partial^2 \mathcal{V}}{\partial \bar{x}_u \partial \bar{x}_v} (x_u - \bar{x}_u) (x_v - \bar{x}_v). \quad (31)$$

$\partial^2 \mathcal{V} / \partial \bar{x}_r^2$ is, in this connexion, called the force constant of the r th bond and $\partial^2 \mathcal{V} / \partial \bar{x}_r \partial \bar{x}_s$ the interaction constant of bonds r and s . In what follows we shall be concerned only with these terms of (31), since the significance of the $\partial^2 \mathcal{V} / \partial \bar{x}_u \partial \bar{x}_v$ (u or $v > n$) cannot be revealed by a consideration of the wave functions of the unsaturation electrons alone.

We shall find it convenient to subdivide $\mathcal{V}$, such that

$$\delta \mathcal{V} = \delta \mathcal{F} + \delta \mathcal{E} \quad (32)$$

for any small change of configuration; $\delta \mathcal{E}$ is the increment in the mobile exchange energy corresponding to this variation in the carbon skeleton of some unsaturated hydrocarbon. $\mathcal{F}$ may be regarded as the contribution to the total energy of the molecule from a 'single-bonded structure', the spins of each of whose electrons are independently quantized. Accordingly, if we neglect changes in energy due to angular distortion and interactions between different σ bonds, we may follow Lennard-Jones (1937), putting

$$\mathcal{F} = \text{const.} + \sum_r \frac{1}{2} \sigma (x_r - s)^2. \quad (33)$$

The term involving the summation over all bonds represents the work required to deform the succession of normal single bonds to attain the values $\{x_r\}$ for the particular configuration under consideration; here σ and s are the force constants and bond length of a pure σ bond respectively. It is tacitly assumed that the different states of hybridization do not appreciably affect our identification of σ and s for unsaturated systems with those measured for, say, ethane.

It only remains, before pursuing our analysis, to express the dependence of the exchange integral J_r on the bond length x_r . Let us suppose that we may represent

the potential energies, as functions of the carbon-carbon internuclear distance x , for ethane and ethylene by the parabolas

$$\text{const.} + \frac{1}{2}\sigma(x-s)^2, \quad \text{const.} + \frac{1}{2}\kappa(x-d)^2,$$

respectively. Then, neglecting complications due to hyperconjugation and hybridization, the non-constant difference between these two quantities may be equated to the non-constant part of the mobile exchange energy for ethylene—that is, $(\frac{2}{3})J$. (For the coefficients of J in the electronic energies of ethane and ethylene are $(-\frac{1}{2})$ and $(-\frac{1}{2} + \frac{2}{3})$ respectively.) Accordingly

$$\frac{2}{3}J_r = \text{const.} + \frac{1}{2}\kappa(x-d)^2 - \frac{1}{2}\sigma(x-s)^2. \quad (34)$$

κ and d are respectively the force constant and bond length in ethylene. We thus have

$$\frac{dJ_r}{dx_r} = \frac{2}{3}\kappa(x_r-d) - \frac{2}{3}\sigma(x_r-s). \quad (35)$$

It seems reasonable to suppose that anharmonicity terms will to some extent cancel out and extend the validity of (35) to cover the entire range $d \leq x \leq s$.

BOND LENGTHS IN EQUILIBRIUM

We have now established a close formal similarity, which will appear shortly, between the pairing or valence-bond approximation on the one hand and the molecular orbital method of Lennard-Jones (1937), Coulson (1939) and Coulson & Longuet-Higgins (1948) on the other.

The condition that the bond lengths x_r assume their values $\bar{x}_r$ is simply that

$$\frac{\partial \mathcal{V}}{\partial x_r} = 0 \quad (r=1, \dots, n).$$

But by (32) and (13)
$$\frac{\partial \mathcal{V}}{\partial x_r} = \frac{\partial \mathcal{F}}{\partial x_r} + \frac{\partial \mathcal{E}}{\partial J_r} \frac{dJ_r}{dx_r} = \frac{\partial \mathcal{F}}{\partial x_r} + \frac{2}{3}p_r \frac{dJ_r}{dx_r}, \quad (36)$$

so that, from (33) and (35), we find that

$$\bar{x}_r = \frac{d\kappa p_r + s\sigma(1-p_r)}{\kappa p_r + \sigma(1-p_r)}. \quad (37)$$

This equation, connecting the length of a particular bond with its order, for HLSP theory, is identical in form with that deduced by means of analogous assumptions by Coulson for the molecular orbital theory. (To complete this analogy, we should, of course, also have the relation

$$3J = 4\beta$$

between the fundamental exchange and resonance integrals respectively of the two theories. For only then would equal bond orders by both methods imply equal resonance energies. This point will not, however, be further elaborated here since we are primarily concerned with variations in bond lengths.)

It is interesting to compare the two methods by means of equation (37). Using the values

$$\left. \begin{aligned} s &= 1.54, & d &= 1.33 \text{ \AA}, \\ \sigma &= 4.96, & \kappa &= 9.80 \times 10^5 \text{ dynes/cm.}, \end{aligned} \right\} \quad (38)$$

we may predict bond lengths for benzene and graphite using calculated values for the bond orders. These have been tabulated in table 1, together with the experimental values.

TABLE 1

molecule	$p(\text{HLSP})$	$x(\text{HLSP})$	$p(\text{MO})$	$x(\text{MO})$	exp. (Å)
benzene	0.623	1.380	0.667	1.372	1.39
graphite	0.45	1.410	0.53	1.395	1.42

We see that the pairing approximation (HLSP) gives perhaps rather better results than the molecular orbital (MO) method. But the agreement with experiment is good in both cases, in view of the approximations which it has been found necessary to introduce. If we were to take $d = 1.34 \text{ \AA}$, as now seems more likely, this agreement would be improved for both theories. We have, however, retained $1.33 \text{ \AA}$ in order not to jeopardize a later comparison of the two methods.

In practice, of course, neither in HLSP nor in MO theory do we use the relation (37) in order to predict bond lengths, but, more reliably, construct curves by plotting values of x as ordinates against values of p as abscissae, using the sets of values (38) together with those for benzene and graphite.

FORCE CONSTANTS AND INTERACTION TERMS

Let us now determine the force and interaction constants of (31) in terms of the bond orders and, as will appear, the mutual bond polarizabilities. We have

$$\frac{\partial^2 \mathcal{V}}{\partial \bar{x}_r^2} = \frac{\partial^2 \mathcal{F}}{\partial \bar{x}_r^2} + \frac{\partial^2 \mathcal{E}}{\partial \bar{x}_r^2} = \frac{\partial^2 \mathcal{F}}{\partial \bar{x}_r^2} + \frac{\partial \mathcal{E}}{\partial \bar{x}_r} \frac{d^2 J_r}{d \bar{x}_r^2} + \frac{\partial^2 \mathcal{E}}{\partial J_r^2} \left(\frac{d J_r}{d \bar{x}_r} \right)^2,$$

whence, using (13), (15), (33), (34) and (37), we find

$$\frac{\partial^2 \mathcal{V}}{\partial \bar{x}_r^2} = \sigma(1 - p_r) + \kappa p_r + \frac{2}{3} \left\{ \frac{\kappa \sigma(s - d)}{\sigma(1 - p_r) + \kappa p_r} \right\}^2 \pi_{r,r}. \quad (39)$$

Similarly, it is found that

$$\frac{\partial^2 \mathcal{V}}{\partial \bar{x}_r \partial \bar{x}_s} = \frac{\partial^2 \mathcal{F}}{\partial \bar{x}_r \partial \bar{x}_s} + \frac{\partial^2 \mathcal{E}}{\partial \bar{x}_r \partial \bar{x}_s} = \frac{\partial^2 \mathcal{E}}{\partial J_r \partial J_s} \frac{d J_r}{d \bar{x}_r} \frac{d J_s}{d \bar{x}_s},$$

from which we derive

$$\frac{\partial^2 \mathcal{V}}{\partial \bar{x}_r \partial \bar{x}_s} = \frac{2}{3} \left\{ \frac{\kappa \sigma(s - d)}{\sigma(1 - p_r) + \kappa p_r} \right\} \left\{ \frac{\kappa \sigma(s - d)}{\sigma(1 - p_s) + \kappa p_s} \right\} \pi_{r,s}. \quad (40)$$

These equations differ from those obtained by Coulson & Longuet-Higgins only in the factor $\frac{2}{3}$, which appears as $\frac{1}{2}$ in the molecular orbital method. The origin of this apparent difference is of course clear when we remember that the $\pi_{r,s}$ are expressed as reciprocals of J and β respectively in the two approaches. Once more the equations

of the HLSP and MO theories lead to formally identical results; that the force constant of some bond r of a conjugated system consists of two parts; the former represents what may be called the 'static' force constant and depends only on the bond order p_r , whereas the latter serves to diminish the 'static' term and represents the resilience of the electronic system in accommodating itself to changes in the $\bar{x}_r$. And that the interaction between two bonds r and s depends both on the bond orders p_r and p_s and on their mutual polarizability $\pi_{r,s}$.

We may apply the formulae (39) and (40) to determine the force and interaction constants in benzene. This example is particularly favourable since our initial assumption of equal J 's is indeed no assumption at all here and, further, in five of the planar vibrations of benzene the bond angles remain unchanged. The mutual bond polarizabilities for the ground state of the benzene molecule take the values

$$\pi_{1,1} = 0.344/J, \quad \pi_{1,2} = -0.304/J, \quad \pi_{1,3} = 0.224/J, \quad \pi_{1,4} = -0.184/J,$$

where the bonds have been numbered in order around the ring. The bond orders are all 0.623, so that, in the notation adopted by Coulson & Longuet-Higgins, the various force and interaction constants are

$$F = 5.95, \quad F_o = 1.79, \quad F_m = -1.32, \quad F_p = 1.08 \times 10^5 \text{ dynes/cm.}$$

To facilitate the comparison of these values with those derived by the MO method, we have taken the value

$$\frac{3}{2}J = 40 \text{ kcal./mole} \quad (41)$$

from measured resonance energies.

These figures may be compared with the empirical values cited by Coulson & Longuet-Higgins, and those which they calculate. It appears that there are only two reliable sets of values which have been determined from the vibrational analysis, namely $(F - F_o - F_m + F_p)$ and $(F + 2F_o + 2F_m + F_p)$. We tabulate our results in table 2. It again appears that the HLSP method is capable of giving, if anything, rather better results than the molecular orbital method. These values do not depend critically upon our choice of values for s , d , σ and κ , so that we may predict force constants with some confidence. If certain ionic structures were to be included in the analysis, the appropriate constants for certain excited states of benzene could also be predicted—and a knowledge of these might facilitate the spectroscopic identification of such states.

TABLE 2

quantity	obs.	HLSP	MO
$F - F_o - F_m + F_p$	5.78	6.56	$6.95 \times 10^5 \text{ dynes/cm.}$
$F + 2F_o + 2F_m + F_p$	7.57	7.97	$8.19 \times 10^5 \text{ dynes/cm.}$

THE INTERACTION OF TWO CONJUGATED SYSTEMS

Let us now consider the interaction of two, initially separate, conjugated systems AH and HB which are joined to form a composite system C , by the elimination of two hydrogen atoms and the formation of a bond from an atom a of A to an atom b of B —the interaction of the two benzene rings across the central bond of diphenyl,

for example. A systematic discussion of this phenomenon has been given by Coulson & Longuet-Higgins using molecular orbital theory, and it is interesting to see that their general results may also be obtained using HLSP theory; the more so since whenever these two very different approximations reach the same or similar conclusions, we may of course develop them with greater confidence.

From the thermochemistry of the systems AB , AB and C , an experimental quantity $R(\text{obs.})$ may be defined by means of the relation

$$R(\text{obs.}) = Q_f(AB) - Q_f(AH) - Q_f(HB) - D(a-b) + D(a-H) + D(H-b), \quad (42)$$

where the $Q_f(\dots)$ are molecular heats of formation and the $D(\dots)$ are bond energies. (Since the $D(\dots)$ are not strictly thermodynamic quantities, but depend on the assumption of additivity for the energies of σ bonds, neither are the $R(\text{obs.})$.) These empirical 'conjugation energies' $R(\text{obs.})$ may, for HLSP theory, be compared with a theoretical quantity $R(\text{calc.})$ defined by

$$R(\text{calc.}) = -\mathcal{E}(AB) + \mathcal{E}(AH) + \mathcal{E}(HB), \quad (43)$$

where the $\mathcal{E}$'s represent the mobile exchange energies of the various hydrocarbons. This definition is easily shown to be appropriate in as much as $\mathcal{E}(AH) + \mathcal{E}(HB)$ represents the mobile exchange energy of a hypothetical molecule AB , in which the spin-inclinations of the unsaturation electrons of the fragment A or B are quantized independently of those of B or A respectively. Clearly R represents the degree of delocalization of the separate fragments in the composite system.

The following theorems may now be proved:

- (i) Conjugation energies are always positive.
- (ii) The order of the bond joining two conjugated systems is always positive.

The first result explains the observed increase in stability which, *ceteris paribus*, is associated thermochemically with such conjugation. The second, in conjunction with (37), shows why the formal single bond connecting two resonating residues always appears partially double-bonded when its length is determined. Further, we may deduce that every bond of a complex conjugating system is necessarily double-bonded in part, unless it is a member of a cyclic system of bonds. These results are quite general, but need modification if steric factors preclude the coplanarity of the two conjugated fragments.

The proof of these theorems is elementary, and follows analogous lines to those put forward for identical theorems by Coulson & Longuet-Higgins in molecular orbital theory. We may imagine the systems A and B to be initially isolated and then to approach so that the atomic orbital ϕ_a of the former overlaps atomic orbital ϕ_b of the latter. Since the relative orientation of the spins of the electrons associated with a and b was originally random, with $\mathbf{s}_a \cdot \mathbf{s}_b = 0$, p_r must have vanished initially. Accordingly, by setting $s = r$ in (14) and integrating, we find

$$p_r = \int_0^{J_r} \pi_{r,r} dJ_r. \quad (44)$$

But, by (17), $\pi_{r,r}$ is always negative and J_r decreases steadily, so that when the systems come together

$$p_r > 0. \quad (45)$$

This proves (ii). Similarly, since by definition

$$R = - \int_0^{J_r} \frac{\partial \mathcal{E}}{\partial J_r} dJ_r, \quad (46)$$

and
$$\frac{\partial \mathcal{E}}{\partial J_r} = \frac{3}{2} p_r > 0,$$

from (45), theorem (i) follows, namely, that

$$R > 0. \quad (47)$$

RESIDUAL AFFINITIES AND CONJUGATING POWER

It is possible, in molecular orbital formalism, to give approximate expressions for R and p_r in terms of the self-polarizabilities of atoms a and b and to show that the conjugating power of, say, an aryl residue is directly related to the self-polarizability of that atom of the residue which is a constituent of the bond permitting its interaction with some other residue. It is difficult to see how such simplifications can be developed for HLSP theory, since self-atom-polarizabilities can only be introduced together with a representative selection of ionic structures. We may, however, investigate the implications of the following assumption; the conjugating power at any point of an unsaturated system may be associated with the residual affinity at that point or carbon atom.

The residual affinity r_i at any atom i of an unsaturated hydrocarbon may be defined (Moffitt 1949) by the relation

$$r_i = \left\{ 1 - \sum_s^{(i)} p_s \right\} + \mathcal{N}_c, \quad (48)$$

where $\sum_s^{(i)}$ sums over all bonds s terminating at i . The term $\mathcal{N}_c$, which ensures a non-negative r_i , assumes the values 0, $\frac{1}{3}$ or $\frac{2}{3}$ according as i has one, two or three nearest neighbours—when we speak of i as belonging to class (i), (ii) or (iii) respectively. These residual affinities have been shown to represent the chemical reactivity at the various carbon atoms of such conjugated molecules and to be inversely correlated with the strength of the π -electronic binding there.

TABLE 3

class (ii) atoms	r.a. (HLSP)	c.p. (MO)	class (i) atoms	r.a. (HLSP)	c.p. (MO)
2-hexatrienyl	0.076	0.389			
phenyl	0.087	0.398			
2-butadienyl	0.089	0.402			
β -naphthyl	0.113	0.405	vinyl	0.000	0.500
2-anthracyl	0.127	0.411	1-butadienyl	0.089	0.626
α -naphthyl	0.127	0.443	1-hexatrienyl	0.117	0.685
1-anthracyl	0.142	0.454			
3-hexatrienyl	0.153	0.470			
10-anthracyl	0.159	0.526			

r.a. = residual affinity; c.p. = conjugating power.

Let us therefore compare the molecular orbital conjugating powers—which Coulson & Longuet-Higgins have successfully compared with observed values of the R 's and p_r 's—with our residual affinities. We draw up table 3 for atoms of classes (i) and (ii), for class (iii) atoms clearly cannot conjugate in this way. In this we compare our residual affinities (r.a.) with the conjugating powers (c.p.) of molecular orbital theory. We see that any sequence of residues with monotonically increasing conjugating powers has, apparently without exception, the additional property of monotonically increasing residual affinities. Our tentative assumption is therefore justified: we may, indeed, associate the conjugating power of a hydrocarbon residue with the residual affinity of the atom which is to initiate conjugation with another residue.

The work described in this paper arose out of the programme of fundamental research undertaken by the Board of the British Rubber Producers' Research Association.

REFERENCES

- Coulson, C. A. 1939 *Proc. Roy. Soc. A*, **169**, 413.
Coulson, C. A. & Longuet-Higgins, H. C. 1948 *Proc. Roy. Soc. A*, **193**, 447; **A, **195**, 188.
Lennard-Jones, J. E. 1937 *Proc. Roy. Soc. A*, **158**, 280.
Moffitt, W. E. 1949 *Trans. Faraday Soc.* **45**, 373.
Pauling, L. 1933 *J. Chem. Phys.* **1**, 280.
Penney, W. G. 1937 *Proc. Roy. Soc. A*, **158**, 306.
Penney, W. G. & Lloyd, E. H. 1939 *Trans. Faraday Soc.* **35**, 835.**

The kinetics of the polymerization of sarcosine carbonic anhydride

BY S. G. WALEY AND J. WATSON

Courtaulds Ltd., Maidenhead, Berks

(Communicated by A. H. Wilson, F.R.S.—Received 17 June 1949)

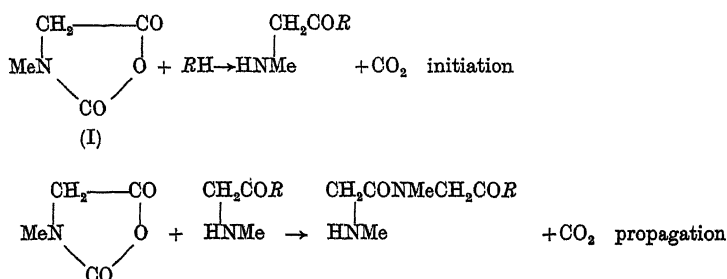
The polymerization of carbonic anhydrides, a reaction important both as a preparative method for the synthesis of polypeptides and as an example of a less familiar type of polymerization, has been studied in detail using sarcosine carbonic anhydride (I). The propagation reaction has been shown to involve a reversibly formed compound between the carbonic anhydride and the polymer; this compound decomposes by both a unimolecular and a bimolecular route. The equilibrium constant for the formation of the intermediate compound, and both velocity constants for its decomposition have been determined in two solvents, and the energies of activation and frequency factors calculated. The molecular weight distribution has been calculated; it is extremely sharp. The viscosities of the polymers prepared by this reaction have been measured, and the relation between the intrinsic viscosity and the molecular weight established.

INTRODUCTION

Polymerizations are customarily divided into two classes: condensation and addition polymerizations. There is, however, a third type of polymerization, several examples of which have been listed by Flory (1940). There has hitherto been no

detailed investigation of the kinetics of this third type of polymerization, but the main features have been elucidated in the polymerization of ethylene oxide (Perry & Hibbert 1940). Polymers of widely varying size were obtained which all reacted with ethylene oxide, and polymerization only stopped when all the ethylene oxide was used up; it could be started again by adding more ethylene oxide. The degree of polymerization was controlled by the relative concentrations of ethylene oxide and initiator and by the progress of reaction (in contradistinction to addition polymerizations). A detailed investigation of an example of this third type of polymerization has now been carried out, this example being the polymerization of sarcosine carbonic anhydride.

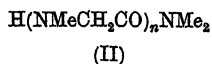
Sarcosine carbonic anhydride (I) reacts with alcohols, water and amines to give an amino-acid derivative and carbon dioxide; this derivative, being an amine, reacts with the carbonic anhydride and thereby generates a new amino group:



The polymers, which are polypeptides, do not react with each other, and the reaction stops only when there is no more carbonic anhydride left. When the polymerization is initiated by lower alcohols, however, the situation may be complicated by the subsequent condensation between terminal amino groups and ester groups. This polymerization of carbonic anhydrides is well known, and has been used, especially recently, as a method for preparing polypeptides (e.g. Astbury, Dalglish, Darmon & Sutherland 1948). There has, however, hitherto been no detailed study of the kinetic mechanism of this reaction. Apart from its intrinsic interest, a knowledge of the kinetic mechanism enables the molecular weight distribution to be predicted, and is necessary to elucidate the composition of co-polymers.

EXPERIMENTAL

Sarcosine carbonic anhydride was chosen for this investigation because it gives polymers soluble in organic liquids, intermolecular hydrogen bonds not being formed to an appreciable extent. The kinetics are simplified by using preformed polymer to initiate the polymerization, so that the propagation step is isolated. In most of the experiments, polysarcosine dimethylamide (II; $n \simeq 10$) was used as initiator. The polymerization was studied in both acetophenone and nitrobenzene solutions.



(a) *Preparation of materials*

Sarcosine carbonic anhydride (I) was prepared by the method of Sigmund & Wessely (1926) and purified first by recrystallization from dry benzene and then by sublimation in a vacuum ($60^{\circ}\text{C}/0.02\text{ mm.}$). It was stored in a vacuum desiccator over phosphorus pentoxide, and was freshly sublimed before use.

Sarcosine dimethylamide (II; $n = 1$) was first prepared by adding the carbonic anhydride to a large excess of dimethylamine; after the dimethylamine had evaporated, the amide was isolated by distillation as a colourless oil; b.p. $64^{\circ}\text{C}/1\text{ mm.}$ (Found C, 51.9; H, 10.6; N, 24.2. $\text{C}_5\text{H}_{12}\text{ON}_2$ requires C, 51.75; H, 10.35; N, 24.2 %). The picrate, recrystallized from ethanol, had m.p. 193 to 195°C (Found C, 38.6; H, 4.5; N, 20.0. $\text{C}_{11}\text{H}_{15}\text{O}_8\text{N}_5$ requires C, 38.25; H, 4.35; N, 20.3 %). The sarcosine dimethylamide was more conveniently prepared in bulk by gradually adding chloracetyl dimethylamide (32 g.) (prepared from chloracetyl chloride and dimethylamine in ether; cf. Freudenberg & Markert 1927) to 33 % w/v ethanolic methylamine (200 ml.). After 5 days the solution was evaporated and the residue basified with 5N-sodium hydroxide, saturated with potassium carbonate, and extracted with benzene. The dimethylamide (20 g.) was isolated by distillation.

N-Isopropylglycine carbonic anhydride was prepared from the amino-acid in the usual way (Found C, 50.0; H, 6.55; N, 10.3. $\text{C}_6\text{H}_9\text{O}_3\text{N}$ requires C, 50.4; H, 6.3; N, 9.8 %). On reaction with dimethylamine it gave *N*-isopropylglycine dimethylamide, b.p. $51^{\circ}\text{C}/0.1\text{ mm.}$ (Found C, 57.9; H, 11.4. $\text{C}_7\text{H}_{16}\text{ON}_2$ requires C, 58.4; H, 11.1 %), characterized by its picrate, which crystallized from ethanol in yellow needles, m.p. 142 to 145°C (Found C, 42.0; H, 5.2; N, 19.1. $\text{C}_{13}\text{H}_{19}\text{O}_8\text{N}_5$ requires C, 41.85; H, 5.1; N, 18.8 %).

Acetophenone was shaken with phosphorus pentoxide, decanted on to further phosphorus pentoxide, and after several hours poured on to anhydrous potassium carbonate and finally fractionated at 0.5 mm.

Nitrobenzene was stood over phosphorus pentoxide overnight, decanted and distilled at 0.05 mm. The same cycle of operations was repeated using first potassium carbonate and then barium oxide; the solvent was finally fractionated, b.p. $40^{\circ}\text{C}/0.02\text{ mm.}$

Benzyl alcohol was distilled over sodium and then redistilled, both distillations being conducted under reduced pressure.

(b) *Measurement of rates*

The rate was followed manometrically by the evolution of carbon dioxide; since the reaction was carried out at pressures below 15 cm. of mercury, the pressure could be taken as proportional to the concentration of carbon dioxide formed. The apparatus used was of the constant-volume type described by Bell (Baughan & Bell 1937; Bell & Trotman-Dickenson 1949), modified by incorporating a glass spiral between the reaction vessel and the manometer. The carbonic anhydride was placed in the reaction vessel, the initiator solution added, and the system evacuated and sealed off promptly. The reaction vessel was agitated in a thermostat by banging, as described by Bell. The reaction may be followed from the start by having the carbonic

anhydride in a small glass bucket loosely suspended on a platinum wire; the agitation tips the bucket into the solution.

The concentration of amino groups at the end of the reaction was measured by diluting the solution with ethanol and titrating against standard acid; it was thus confirmed that the concentration of amino groups $[X]_0$ was the same as at the beginning of the reaction. In agreement with this, by the addition of further carbonic anhydride at the end of the reaction, polymerization recurs at a rate consonant with the original value of $[X]_0$.

RESULTS

When $\log(p_\infty - p)$ was plotted against the time, the curves were of the form shown in figure 1 (where p is the pressure of carbon dioxide in mm. Hg at time t , and p_∞ the final pressure). This shape indicates that the reaction starts with an apparent order

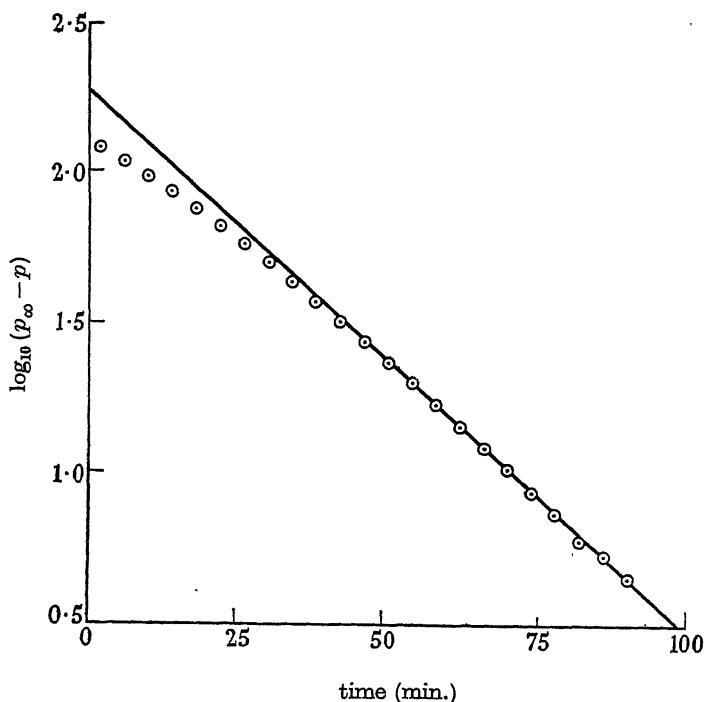


FIGURE 1. Solvent, nitrobenzene; temperature 25°C ; $[M]_0 = 0.123$ mol./l. and $[X]_0 = 0.0052$ mol./l.

between zero and 1, and approaches first order as the reaction proceeds. The gradient k_{obs} of the straight-line portion is referred to as the end-gradient. This end-gradient was independent of $[M]_0$ (where $[M]_0$ is the initial concentration of carbonic anhydride), and was the same in duplicate runs carried out by adding more anhydride at the end of the first run. Thus the reactivity of the polymer is apparently independent of chain length over the range covered by these experiments.

(a) Experiments on acetophenone solution

(1) Sarcosine carbonic anhydride (0.5 g.) was added to an approximately 0.02N solution of dry dimethylamine in acetophenone (25 ml.). After the reaction had ended any unused dimethylamine was removed by a stream of nitrogen and the solution diluted to 40 ml. The concentration of amino groups was found to be

TABLE 1				
25° C			40° C	
$10^3 [X]_0$ (mol./l.)	$10^4 k_{\text{obs.}}$ (sec. ⁻¹)	$k_{\text{obs.}}/[X]_0$ (l.mol. ⁻¹ sec. ⁻¹)	$10^4 k_{\text{obs.}}$ (sec. ⁻¹)	$k_{\text{obs.}}/[X]_0$ (l.mol. ⁻¹ sec. ⁻¹)
12.75	33.55	0.263	30.0	0.235
	32.8	0.258	30.5	0.239
	31.2	0.245	—	—
8.5	15.2	0.179	15.7	0.185
	16.8	0.197	16.2	0.191
	—	—	18.4	0.217
6.375	8.71	0.137	—	—
	7.75	0.122	—	—
	9.15	0.143	—	—
4.25	4.70	0.111	6.25	0.147
	4.30	0.100	6.05	0.142
1.16	0.785	0.067	1.13	0.0975
	—	—	1.16	0.100

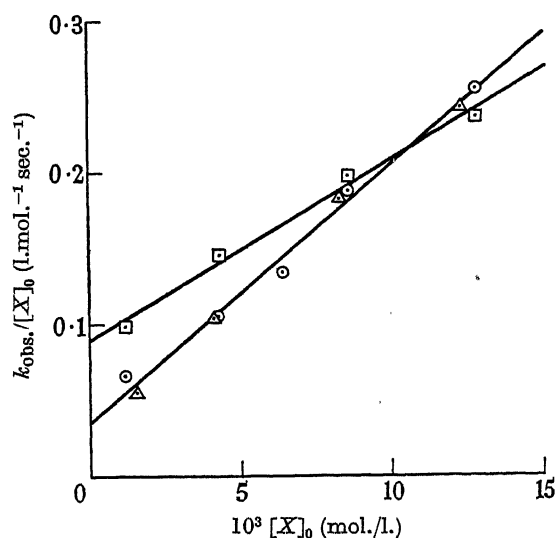


FIGURE 2. Solvent, acetophenone. $\odot$ Solution 1 at 25° C, $\triangle$ solution 2 at 25° C and $\square$ solution 1 at 40° C.

0.01275 mol./l. by titration in alcoholic solution against aqueous N/100-hydrochloric acid, using 2-(4'-benzylideneaminophenyl)-quinoline as indicator. This substance, whose preparation has already been described (Waley 1948), is a convenient indicator for titrating weak bases in organic solvents. The results obtained with this polymer solution are given in table 1, and in figure 2 $k_{\text{obs.}}/[X]_0$ is plotted against $10^3 [X]_0$.

The linear relation expressed by figure 2 shows that we may write

$$k_{\text{obs.}} = a'[X]_0 + b'[X]_0^2. \quad (1)$$

The figure also shows that the parameter b' has a negative temperature coefficient.

(2) Sarcosine carbonic anhydride (2g.) was added to a 0.04N solution of dimethylamine in dry dioxan (70 ml.). At the end of the reaction, precipitation was completed by the addition of dry ether, the solid polymer collected and washed with ether at the centrifuge and dried at 40 to 60° C/0.05 mm. It was then dissolved in acetophenone (100 ml.), giving a solution of $[X]_0 = 0.0123$ mol./l. The results obtained with this solution at 25° C (shown in table 2 and figure 2) agree with those obtained with the polymer solution described above.

TABLE 2. 25° C

$10^3 [X]_0$ (mol./l.)	$10^4 k_{\text{obs.}}$ (sec. ⁻¹)	$k_{\text{obs.}}/[X]_0$ (l.mol. ⁻¹ sec. ⁻¹)
12.3	29.25	0.238
	30.45	0.248
8.22	15.0	0.182
	15.1	0.184
4.11	4.33	0.105
	4.34	0.105
1.54	0.86	0.056

(b) *Experiments in nitrobenzene solution*

Sarcosine carbonic anhydride (0.75 g.) was decomposed with sarcosine dimethylamide (0.09 ml.) in nitrobenzene (50 ml.). This solution had $[X]_0 = 0.0155$ mol./l. and was used to obtain the results given in table 3 and figure 3; the results are similar to those obtained in acetophenone solution.

TABLE 3

$10^3 [X]_0$ (mol./l.)	25° C		50° C	
	$10^4 k_{\text{obs.}}$ (sec. ⁻¹)	$k_{\text{obs.}}/[X]_0$ (l.mol. ⁻¹ sec. ⁻¹)	$10^4 k_{\text{obs.}}$ (sec. ⁻¹)	$k_{\text{obs.}}/[X]_0$ (l.mol. ⁻¹ sec. ⁻¹)
15.5	41.5	0.275	36.2	0.233
	42.6	0.282	40.7	0.263
	60.8	0.392	40.4	0.266
	39.4	0.254	38.3	0.247
10.3	25.5	0.247	21.65	0.210
	24.4	0.237	24.2	0.235
	18.9	0.184	15.5	0.151
	20.1	0.195	26.6	0.258
5.17	6.90	0.133	8.06	0.156
	6.80	0.131	9.99	0.193
	7.16	0.139	9.21	0.178
	6.63	0.128	9.18	0.178
3.1	2.61	0.084	3.22	0.104
	4.46	0.144	4.22	0.136
	3.85	0.124	3.90	0.126
	3.37	0.092	5.06	0.164

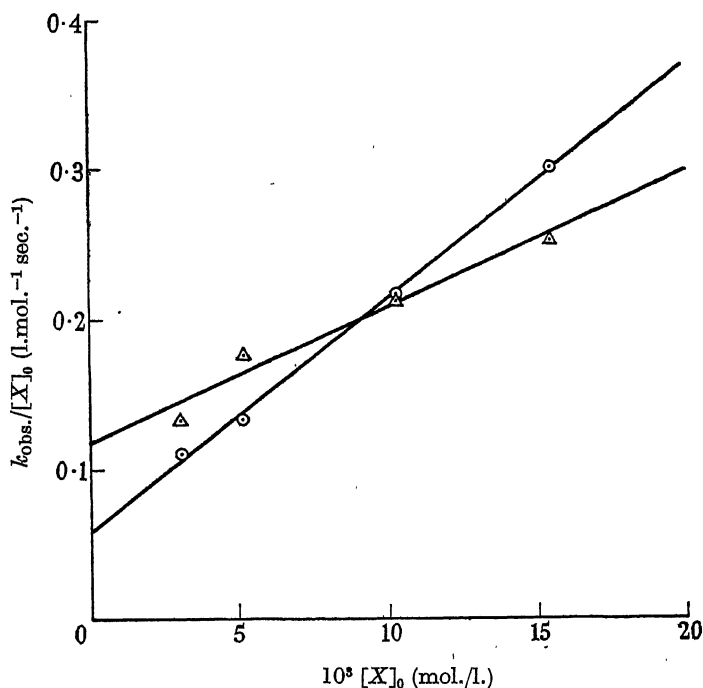


FIGURE 3. Solvent, nitrobenzene. ○ 25° C, △ 50° C.

(c) *Experiments with benzyl alcohol*

A few experiments were carried out on polymerizations initiated by an alcohol. Benzyl alcohol was chosen for this purpose, since it is not too volatile, and because the bulky benzyl ester group is unlikely to react with the terminal amine groups. The main feature of those runs was a very long induction period; thus figure 4 (circles) shows a run in acetophenone solution at 25° C in which $[M]_0 = 0.1461$ mol./l. and the initial concentration of benzyl alcohol ($[A]_0$) = 0.5834 mol./l. This induction period was markedly reduced in the presence of added polymer; thus figure 4 (triangles) shows another run in acetophenone solution at 25° C in which $[M]_0 = 0.1387$ mol./l., $[A]_0 = 0.5318$ and $[X]_0 = 9.76 \times 10^{-5}$ mol./l.

(d) *Viscosity measurements*

A series of polysarcosine dimethylamides was prepared from sarcosine carbonic anhydride and sarcosine dimethylamide in nitrobenzene. Although the solutions were homogeneous throughout the reaction, the higher polymers separated as gels on attempting to filter the solution. At the end of the reaction, therefore, the nitrobenzene solution was diluted with petrol and extracted with water. The aqueous solution was washed with petrol and filtered.

The concentration of polymer was determined by a Kjeldahl nitrogen estimation, which also showed that no appreciable losses occurred during the extraction; the concentration of amino groups was determined by titration in aqueous alcoholic solution using the benzylideneaminophenylquinoline as indicator. The degree of

polymerization thus obtained agreed closely with the ratio of the initial concentration of carbonic anhydride to that of the sarcosine dimethylamide.

The viscosities of the aqueous solutions were measured at 25° C in a B.S.S. no. 1 viscometer. The specific viscosity of each polymer solution was measured at several concentrations, and the intrinsic viscosity $[\eta]$ was found by extrapolating to infinite dilution the plot of $\eta_{sp.}/c$ against c , where c is the base molar concentration. The results are shown in table 4.

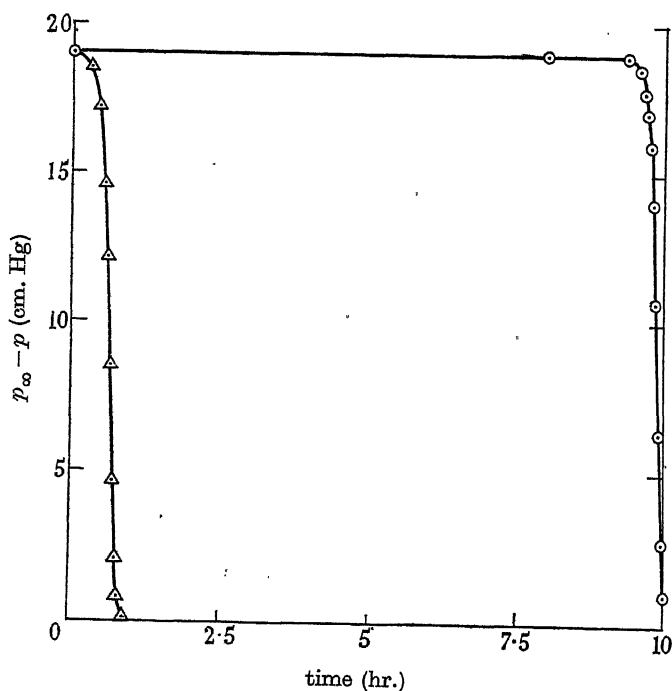


FIGURE 4

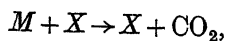
TABLE 4

$[\eta]$	molecular weight
0.356	861
0.478	1,825
0.680	3,524
0.870	5,260
1.143	8,200
1.384	13,280
1.745	22,040

DISCUSSION

(a) *Kinetics and mechanism of reaction*

The simplest kinetic scheme for the polymerization of sarcosine carbonic anhydride initiated by preformed polymer is



where M denotes the carbonic anhydride and X any polymer species. This scheme involves only one velocity constant k , that of propagation. On this basis, the molar concentration $[X]_0$ of polymer is constant, and the rate of disappearance of carbonic anhydride would be given by

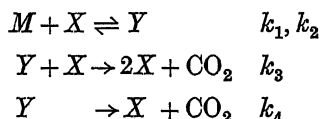
$$-d[M]/dt = k[M][X]_0 = d[CO_2]/dt, \quad (2)$$

assuming that k is independent of chain length. If this equation were valid the pressure-time relationship would be

$$\log(p_\infty - p) = k[X]_0 t. \quad (3)$$

The experimental plots of figure 1, however, were of the linear form implied by this equation only towards the end of the reaction. Equation (3) also implies that on comparing runs with different values of $[X]_0$ the gradient of the plot of $\log(p_\infty - p)$ against t should be a linear function of $[X]_0$. In fact, it is $k_{\text{obs.}}/[X]_0$, and not the end-gradient $k_{\text{obs.}}$, that varies linearly with $[X]_0$ (equation (1)). Furthermore, for some values of $[X]_0$, $k_{\text{obs.}}$ had a negative temperature coefficient. This implies the inclusion of an equilibrium constant in the kinetic scheme.

The following kinetic scheme accounts for the main features mentioned above, and fits the experimental curves obtained:



where Y is a complex formed from the anhydride and any amino group. Thus

$$-d[M]/dt = k_1[M][X] - k_2[Y], \quad (4)$$

$$d[Y]/dt = k_1[M][X] - k_2[Y] - k_3[X][Y] - k_4[Y]. \quad (5)$$

If the initial concentrations of M and X are $[M]_0$ and $[X]_0$, the following relations are stoichiometrically necessary:

$$[X] + [Y] = [X]_0, \quad (6)$$

$$[CO_2] = [M]_0 - [M] - [Y], \quad (7)$$

where $[CO_2]$ is the total number of moles of carbon dioxide in the system divided by the volume of solution.

If $k_3 = k_4 = 0$, then $[Y] + [M] = \text{constant} = [M]_0$, and it can readily be shown that

$$[M] = \frac{a + be^{-\gamma t}}{1 + ce^{-\gamma t}}, \quad [Y] = [M]_0 - [M] = \frac{f + ge^{-\gamma t}}{1 + ce^{-\gamma t}}. \quad (8)$$

Here a , b , c and γ are constants given by the relations

$$\begin{aligned} a &= \lambda_1, \quad b = -\lambda_2([M]_0 - \lambda_1)/([M]_0 - \lambda_2), \quad c = -([M]_0 - \lambda_1)/([M]_0 - \lambda_2), \\ \gamma &= k_1(\lambda_1 - \lambda_2), \end{aligned} \quad (9)$$

where λ_1 and λ_2 are the positive and negative roots respectively of

$$k_1\lambda^2 + \{k_1([X]_0 - [M]_0) + k_2\}\lambda - k_2[M]_0 = 0. \quad (10)$$

Thus, in a time of order $1/\gamma$, $[M]$ effectively reaches the equilibrium value a .

If k_3 and k_4 are not zero, but small compared with k_1 and k_2 , the solution of the equations can be expressed in the above form with a , b , c , f and g slowly varying functions of the time, i.e. with da/dt small compared with γa , and so on. An approximate solution valid for times much greater than $1/\gamma$, for which $[M]$ is effectively equal to $a(t)$, can be found as follows.

On differentiating equation (4) we find

$$(1 + K[M]) \frac{d[Y]}{dt} = K[X] \frac{d[M]}{dt} + \frac{1}{k_2} \frac{d^2[M]}{dt^2}, \quad (11)$$

and, on rearranging the same equation,

$$[Y] = \frac{K[M][X]_0}{1 + K[M]} + \frac{1}{k_2(1 + K[M])} \frac{d[M]}{dt}, \quad (12)$$

where $K = k_1/k_2$. Hence, substituting for $[Y]$ and $d[Y]/dt$ in equation (5), we have

$$\left(1 + \frac{K[X]}{1 + K[M]}\right) \frac{d[M]}{dt} = -\left(\frac{k_3[X]_0}{1 + K[M]} + k_4\right) \left(\frac{K[X]_0[M]}{1 + K[M]} + \frac{1}{k_2(1 + K[M])} \frac{d[M]}{dt}\right) + \frac{k_3 K[X]_0[M]}{k_2(1 + K[M])^2} \frac{d[M]}{dt} + \frac{k_3}{k_2^2(1 + K[M])^2} \left(\frac{d[M]}{dt}\right)^2 - \frac{1}{k_2(1 + K[M])} \frac{d^2[M]}{dt^2}. \quad (13)$$

It is clear from this equation that, for times large compared with $1/\gamma$,

$$d[M]/dt = O(\alpha[M]), \quad \text{where } \alpha = O\{K[X]_0(k_3[X]_0 + k_4)\}.$$

Also, by differentiating equation (13) we see that $d^2[M]/dt^2 = O(\alpha d[M]/dt)$, and so on. Equation (13) can therefore be solved by successive approximations, the first approximation being given by neglecting terms containing $k_3 d[M]/dt$, $k_4 d[M]/dt$, $(d[M]/dt)^2$ and $d^2[M]/dt^2$. This gives

$$\left(1 + \frac{K[X]_0}{(1 + K[M])^2}\right) \frac{d[M]}{dt} = -\left(\frac{k_3[X]_0}{1 + K[M]} + k_4\right) \frac{K[X]_0[M]}{1 + K[M]}, \quad (14)$$

since, with the present approximations,

$$[X] = [X]_0/(1 + K[M]), \quad [Y] = K[M][X]_0/(1 + K[M]). \quad (15)$$

Equation (14) can be integrated at once to

$$t = A([M]_i - [M]) + B \log \frac{[M]_i}{[M]} + C \log \left(\frac{k_4 K[M]_i + k_4 + k_3[X]_0}{k_4 K[M] + k_4 + k_3[X]_0} \right), \quad (16)$$

where $A = \frac{1}{k_4[X]_0}$, $B = \frac{1 + K[X]_0}{K[X]_0(k_4 + k_3[X]_0)}$, $C = -\frac{k_3^2[X]_0 + k_4^2 K}{k_4^2 K(k_4 + k_3[X]_0)}$,

and $[M]_i$ is the value of $[M]$ at $t = 0$. Since these equations are only valid for times much larger than $1/\gamma$, the initial value of $[M]$ is not $[M]_0$ but the 'equilibrium value' $a = \lambda_1$ given by equations (9) and (10). It can easily be verified that this is the same value as is obtained from equation (7) by putting $[\text{CO}_2] = 0$.

Towards the end of the reaction equation (16) reduces to

$$\log [M] = \text{constant} - t/B. \quad (17)$$

But $\{1 + K[X]_0/(1 + K[M])\}[M] = [M]_0 - [\text{CO}_2]$

(see equations (7) and (15)), and towards the end of the reaction we can write

$$(1 + K[X]_0)[M] = [M]_0 - [\text{CO}_2].$$

Hence, from equation (17),

$$\log ([M]_0 - [\text{CO}_2]) = \text{constant} - t/B,$$

$$\text{i.e.} \quad \log (p_\infty - p) = \text{constant} - t/B. \quad (18)$$

Comparison of equation (18) with the empirical equation (1) shows that $1/B$ is equal to the parameter k_{obs} , provided that $K[X]_0$ is small compared to 1. The results given later agree with this assumption.

The use of equation (18) is limited not only by the approximations involved in its derivation, but even more by the practical difficulty of measuring end-gradients accurately. Since both the parameters a' and b' in equation (1) involve an equilibrium constant their temperature coefficients may be negative.

The experimental plots of $p_\infty - p$ against time were fitted by employing the theoretical equation (16) in the following way. A trial value of the equilibrium constant K was first taken and a value of $[M]_i$ calculated from equation (7). An arbitrary set of values for $[M]$, varying from $[M] = [M]_i$ to $[M] = 0$, was then taken, and using assumed values for k_3 and k_4 values of the time were computed from equation (16). The corresponding values of $[M]$ were then converted into values of $[\text{CO}_2]$ by using equations (7) and (15), thus giving a theoretical $[\text{CO}_2]$, t relation. The calculated values of $[M]_0 - [\text{CO}_2]$, in mol./l., were converted into $p_\infty - p$ in cm. Hg by using an empirical calibration factor such that the first theoretical point lay on the experimental curve. This is necessary to take into account the finite time taken by the carbonic anhydride to dissolve. The values of k_{obs} (the end-gradient) for various values of $[X]_0$ gave approximate values of $k_4 K$ and $k_3 K$, so that for any assumed value of K rough values of k_3 and k_4 were known. Trial and error fitting enables a value of K to be found which is accurate to about 25 %. Table 5 shows that the expression (16) fits the experimental curve accurately.

TABLE 5

Solvent, nitrobenzene; temperature 25° C; $[M]_0 = 0.076$ mol./l.; $[X]_0 = 0.0103$ mol./l. The theoretical points were calculated using $K = 4.0$ l.mol.⁻¹; $k_3 = 4.2$ l.mol.⁻¹ sec.⁻¹; $k_4 = 0.0198$ sec.⁻¹.

time (min.)	$p_\infty - p$ (cm. Hg)	
	obs.	calc.
4.86	8.145	8.145
7.04	6.285	6.29
9.85	4.42	4.42
13.96	2.56	2.56
17.23	1.63	1.62
23.20	0.67	0.685
28.62	0.30	0.31

The values of the constants obtained from the fitting of several experimental curves are given in table 6; table 7 shows the corresponding energies of activation and frequency factors. The uncertainty in the determination of K limits the accuracy with which k_3 and k_4 can be found to about 25 %.

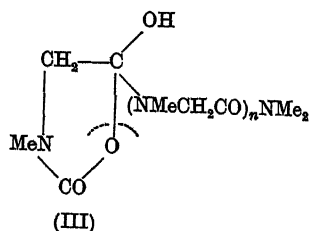
TABLE 6

temp. (° C)	nitrobenzene			acetophenone		
	K (l.mol. ⁻¹)	k_3 (l.mol. ⁻¹ sec. ⁻¹)	k_4 (sec. ⁻¹)	K (l.mol. ⁻¹) ^a	k_3 (l.mol. ⁻¹ sec. ⁻¹)	k_4 (sec. ⁻¹)
25	4.0	4.2	0.0198	3.0	6.0	0.013
40	—	—	—	1.5	9.0	0.060
50	1.4	7.0	0.07	—	—	—

TABLE 7

	nitrobenzene	acetophenone
E_3 (kcal./mol.)	3.9	5.0
A_3 (l.mol. ⁻¹ sec. ⁻¹)	3.1×10^8	2.8×10^4
E_4 (kcal./mol.)	9.6	18.8
A_4 (sec. ⁻¹)	2.3×10^5	1.0×10^{12}
ΔH (kcal./mol.)	-8	-8.5

The complex Y could have a structure of the following kind:



The rate-determining step is envisaged as the rupture of the C—O link marked; the heterolytic fission of this bond will be aided by the approach of another molecule of base, which may act by removing a proton from the oxygen atom. Thus a transition state of the dipolar type, associated with a low-frequency factor, may be involved. In the less facile unimolecular mode of decomposition of Y , the values of the energies of activation and frequency factors are markedly different for the two solvents, suggesting that in this case the solvent plays a part in the scission of the ring.

Direct confirmation of the presence of the complex Y might be provided by examination of the infra-red spectrum during the reaction. At the high concentrations required by this method, however, the polymerization proceeds too fast for measurements to be made. The homologous N -isopropylglycine carbonic anhydride is much less reactive, and the changes occurring in the 3μ region of the infra-red absorption spectrum during the progress of the polymerization initiated by N -isopropylglycine dimethylamide have been examined by Mr E. J. Ambrose (who gives a full report in the appendix). In this region there is a stretching frequency of the N—H bond in the N -isopropylglycine dimethylamide which is markedly modified in the reaction

mixture (figure 5). As the reaction proceeds this modification of the band lessens until finally it is indistinguishable from the band in the pure *N*-isopropylglycine dimethylamide. Thus there is some interaction between the *N*-isopropylglycine carbonic anhydride and the basic nitrogen atom of the amine, and this lends strong support to the hypothesis of the complex (*Y*) in the polymerization of sarcosine carbonic anhydride. The nature of the complex is less certain, and the structure (III) is only one of several possible structures between which the infra-red data do not discriminate.

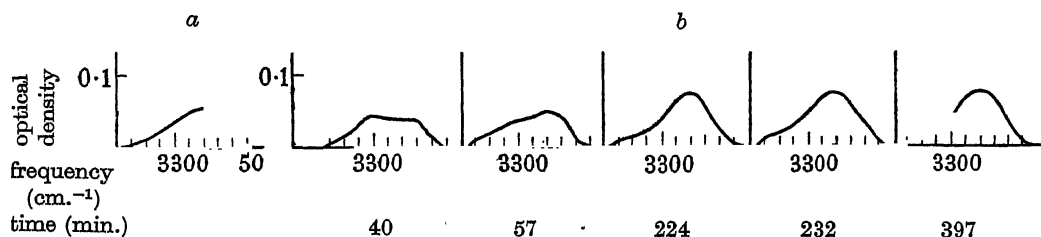
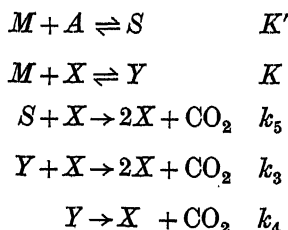


FIGURE 5

The role of the base in the decomposition of sarcosine carbonic anhydride is strikingly shown by the experiments which were carried out using benzyl alcohol as initiator. As described above, there is an induction period which is greatly cut down by the presence of added base, and in this case we may assume that there is a complex *S* formed between the carbonic anhydride and the alcohol, and that this complex only forms polymers at an appreciable rate in the presence of added base. We may thus use the following kinetic scheme:



We assume that both equilibria are rapidly attained, and also that $[S]$ is small. It may readily be shown that

$$-\frac{d[A]}{dt} = K'k_5 \frac{[A][M]}{1 + K[M]} ([A]_0 + [X]_0 - [A]). \quad (19)$$

If we now use the time variable u defined by

$$u = \int_0^t \frac{K[M]}{1 + K[M]} dt, \quad (20)$$

we have

$$A = \frac{[X]_0 + [A]_0}{1 + ([X]_0/[A]_0) e^{K'k_5([A]_0 + [X]_0)u/K}}. \quad (21)$$

By a method similar to that given above it can be shown that

$$\left\{1 + K[M] + \frac{K([A]_0 + [X]_0 - [A])}{1 + K[M]}\right\} \frac{d[M]}{dt} = -KK'k_5[M]^2[A] \frac{([A]_0 + [X]_0 - [A])}{1 + K[M]} - K'k_5[A][M]([A]_0 + [X]_0 - [A]) + Kk_3[M] \frac{([A]_0 + [X]_0 - [A])^2}{1 + K[M]} + Kk_4[M]([A]_0 + [X]_0 - [A]). \quad (22)$$

From equations (20) and (21) this can be written as

$$-k_5 \frac{K'}{K} \left\{ \frac{1+cx}{B_0 c} + \frac{Kx}{(1+K[M])^2} \right\} \frac{d[M]}{dx} = \frac{K'k_5[M]}{(1+K[M])(1+cx)} + \frac{K'k_5}{K(1+cx)} + \frac{k_4}{B_0} + \frac{k_3 cx}{(1+K[M])(1+cx)}, \quad (23)$$

where $B_0 = [A]_0 + [X]_0$, $c = [X]_0/[A]_0$ and $x = e^{K'k_5 B_0 u/K}$.

The validity of this expression is confirmed in the following way. The concentration of amino groups at the end of reaction gives the value of $[A]$ when $[M] = 0$. The finite value u_∞ of u when $[M] = 0$ is obtained by numerical integration of equation (20),

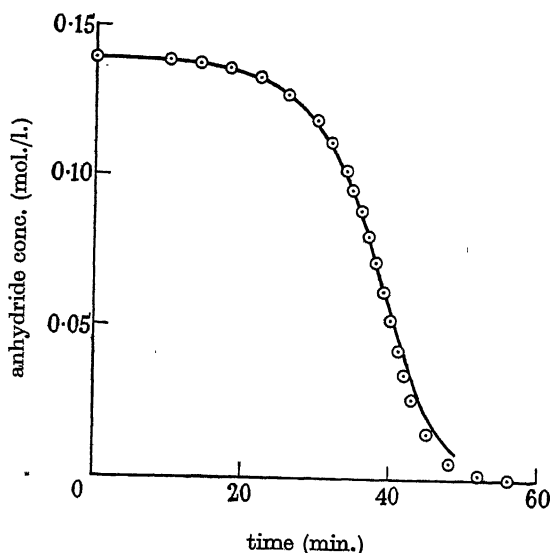


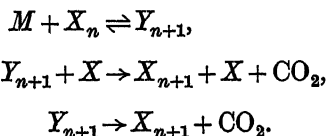
FIGURE 6. Solvent, acetophenone; temperature 25° C; O experimental points. $[M]_0 = 0.139$ mol./l.; $[A]_0 = 0.532$ mol./l.; $[X]_0 = 9.76 \times 10^{-5}$ mol./l. Continuous curve is theoretical calculated from equation (23).

assuming the experimental values of $p_\infty - p$ to be equal to $[M]$ (this is quite a good approximation, since $[Y]$ is small compared with $[M]$ throughout the reaction), and using the previous value of K . These data enable $k_5 K'$ to be obtained from equation (21); $k_5 K'$ is thus found to be $4.23 \times 10^{-2} \text{ l.}^2 \text{ mol.}^{-2} \text{ sec.}^{-1}$. Using the values of K , k_3 and k_4 obtained from polymer-initiated runs in acetophenone at 25°, equation (23) was solved numerically and hence a theoretical $[M]$, t curve drawn by numerical integration of equation (20). This theoretical curve is directly compared with the experimental $([M]_0 - [\text{CO}_2])$, t curve in figure 6. The agreement is, in fact, quite as

good as can be expected, in view of the uncertainty of the exact time origin. The magnitude of the long induction period observed in the absence of added polymer (figure 4) cannot be quantitatively accounted for either by the assumption that traces of polymer contaminate the anhydride or by the assumption that the complex S may decompose by a unimolecular route. The induction period could be due to traces of impurity which only affect the reaction in the absence of added base, even small amounts of which ($[X]_0 \simeq 10^{-4}$) are sufficient to enable the reaction to proceed in a predictable way.

(b) *Molecular weight distributions*

Flory (1940) has shown that the molecular weight distribution derived from the kinetics of the ethylene oxide type of polymerization is much narrower than that obtained from condensation polymerizations. We will now show that the scheme given above for the polymerization of sarcosine carbonic anhydride (initiated by polymer) leads to essentially the same distribution. For this purpose a stage of the reaction may be written



Here X_n refers to the free polymer containing n units, Y_{n+1} the complex formed between this polymer and the carbonic anhydride (i.e. potentially containing $n+1$ units), and X is any free polymer species. If we define $[Z_n]$ by the relation

$$[Z_n] = [X_n] + [Y_{n+1}],$$

we may, as a first approximation, write

$$[Y_{n+1}] = K[M][Z_n]/(1 + K[M]) \quad (24)$$

(compare equation (15)). We will calculate the polymer fraction containing n units, i.e. $[Z_n]/[X]_0$, where $[X]_0$ is the total molar concentration of polymer. We are thus concerned with the change in the total concentration of the n -mer, which is given by

$$\begin{aligned} d[Z_n]/dt &= k_3[Y_n][X] + k_4[Y_n] - k_3[Y_{n+1}][X] - k_4[Y_{n+1}], \\ \text{i.e.} \quad \frac{d[Z_n]}{dt} &= \frac{K[M]}{1 + K[M]} \left(\frac{k_3[X]_0}{1 + K[M]} + k_4 \right) ([Z_{n-1}] - [Z_n]). \end{aligned} \quad (25)$$

If we write this equation as

$$d[Z_n]/dt = \phi([Z_{n-1}] - [Z_n])$$

and define the variable F by $F = \int_0^t \phi dt$,

then $d[Z_n]/dF = [Z_{n-1}] - [Z_n]$,

$$\text{i.e.} \quad \frac{d}{dF}([Z_n] e^F) = e^F [Z_{n-1}] \quad (26)$$

Now $[Z_0] = 0$, and since $F = 0$ and $[Z_1] = [X]_0$ at $t = 0$, we have

$$[Z_1] = [X]_0 e^{-F}.$$

Thus from equation (26)

$$[Z_2] = F[X]_0 e^{-F},$$

and

$$[Z_3] = \frac{1}{2}F^2[X]_0 e^{-F}.$$

Thus

$$\frac{[Z_n]}{[X]_0} = \frac{F^{n-1}}{(n-1)!} e^{-F}. \quad (27)$$

This equation is of the same form as that obtained by Flory.

We have defined ϕ as
$$\phi = \frac{K[M]}{1 + K[M]} \left(\frac{k_3[X]_0}{1 + K[M]} + k_4 \right),$$

i.e.

$$\phi = \frac{[Y]}{[X]_0} (k_3[X] + k_4).$$

Also since

$$d[\text{CO}_2]/dt = k_3[Y][X] + k_4[Y],$$

$$\phi = (1/[X]_0) d[\text{CO}_2]/dt;$$

therefore

$$F = \frac{1}{[X]_0} \int_0^t \frac{d[\text{CO}_2]}{dt} dt,$$

i.e.

$$F = [\text{CO}_2]/[X]_0.$$

Thus the molecular weight distribution at any stage of the reaction can readily be calculated. At the end of the reaction, $[M] = 0$, $[Y] = 0$, and $[X_n] = [Z_n]$, so that

$$\frac{[X_n]}{[X]_0} = \frac{([M]_0/[X]_0)^{n-1}}{(n-1)!} e^{-[M]_0/[X]_0}. \quad (28)$$

When a second polymerization is carried out by adding more carbonic anhydride to the solution obtained from the first, the molecular weight distribution is of the same form, since the sum of two Poisson distributions is itself a Poisson distribution.

In order to compare the sharpness of molecular weight distributions, it is convenient to use the coefficient of variation V as a measure of the relative homogeneity.

The weight fraction of n -mer derived from equation (28) is given by

$$\omega_n = \frac{e^{-\mu}}{1 + \mu} \frac{n\mu^{n-1}}{(n-1)!},$$

where $\mu = [M]_0/[X]_0$. It can be shown that the mean value $\bar{n}$ of n (the weight average degree of polymerization) is given by

$$\bar{n} = \frac{\mu^2 + 3\mu + 1}{\mu + 1},$$

while the standard deviation σ is given by

$$\sigma = (\mu^3 + 2\mu^2 + 2\mu)^{1/2} / (\mu + 1).$$

Thus

$$V = \sigma/\bar{n} = (\mu^3 + 2\mu^2 + 2\mu)^{1/2} / (\mu^2 + 3\mu + 1). \quad (29)$$

As an approximation, for large values of μ ,

$$V \simeq 1/\mu^{\frac{1}{2}}.$$

Clearly, V decreases with increasing molecular weight.

For condensation polymerizations, the weight fraction of n -mer is given by

$$\omega_n = np^{n-1}(1-p)^2,$$

where p is the extent of reaction. The weight average degree of polymerization $\bar{n}$ and the standard deviation σ are given by

$$\bar{n} = (1+p)/(1-p), \quad \sigma = (2p)^{\frac{1}{2}}/(1-p),$$

whence

$$V = (2p)^{\frac{1}{2}}/(1+p). \quad (30)$$

Thus V here increases with increasing molecular weight, and

$$V_{p \rightarrow 1} = 1/\sqrt{2}.$$

Some values for the coefficients of variation for both types of distribution are compared in table 8, which clearly brings out the difference in molecular weight distribution between the two types of polymerization.

TABLE 8

$\bar{n}$	V (from equation (30))	V (from equation (29))
10	0.70	0.29
50	0.71	0.14
100	0.71	0.10
500	0.71	0.045
1000	0.71	0.032

The intrinsic viscosities of several polysarcosine dimethylamides have been measured in water. Since the molecular weight distribution is so sharp, all polymers but the lowest can be considered homogeneous. Thus the parameters C and α in the relationship due to Houwink (1940)

$$[\eta] = C(\text{molecular weight})^\alpha,$$

determined by using the number average molecular weight must be very close to the true values for the homogeneous polymers. The results are given in table 4, and plotted in figure 7, whence the values of C and α are found to be 9.91×10^{-3} and 0.52 respectively. Thus in this case the deviation from the simple Staudinger relation ($\alpha = 1$) is particularly marked.

APPENDIX

The presence of an association compound formed between *N*-isopropylglycine dimethylamide and *N*-isopropylglycine carbonic anhydride in nitrobenzene solution has been demonstrated by observations on the absorption band at 3320 cm.^{-1} , due to the N-H stretching frequency of the amine.

The spectra of the amine and the anhydride were first measured separately in the pure state under high dispersion in the 3μ region using a calcium fluoride prism. In the amine three C-H bands (frequencies 2965, 2932 and 2874 cm^{-1}) and one N-H band (frequency 3316 cm^{-1}) were observed. The anhydride gave only C-H bands at frequencies of 2978, 2939 and 2881 cm^{-1} . It was found that nitrobenzene gave some slight general absorption but no observable bands in the region of 3320 cm^{-1} .

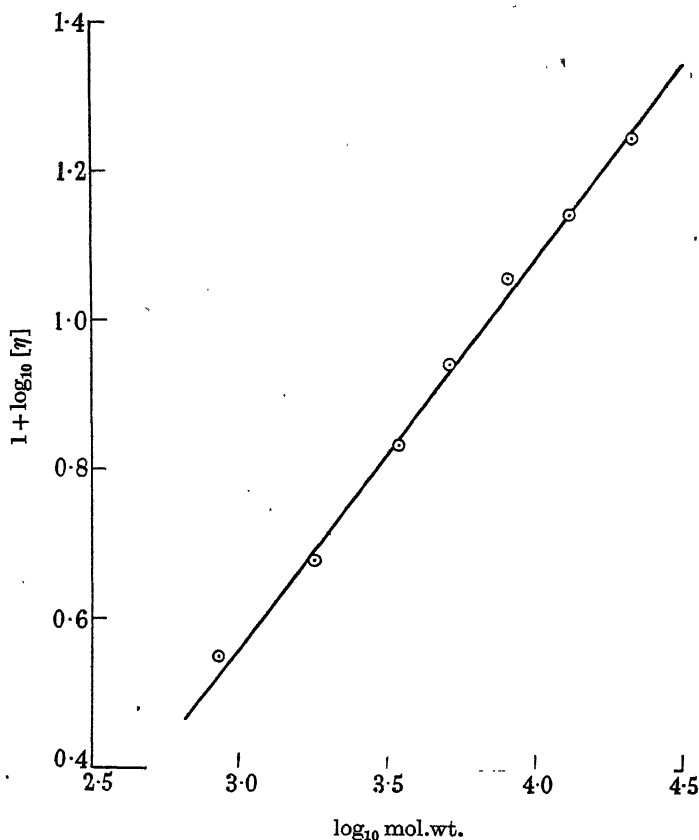


FIGURE 7

The spectrum of a 10 % solution of the amine in nitrobenzene is shown in figure 5*a*. The frequency of the N-H band is 3320 cm^{-1} . An absorption cell was made from two glass cover-slips and mounted on a rolling carriage, which moved the specimen in and out of the beam of radiation automatically, as described by Elliott, Ambrose & Temple (1948). In this way the transmissions of the instrument with and without the liquid film in position were recorded simultaneously and the contours of the absorption bands could be determined accurately.

A reaction mixture was prepared containing 10 % by weight of amine and carbonic anhydride in nitrobenzene, and samples from this mixture were placed in absorption cells after successive intervals of time. Spectra obtained with these samples are shown in figure 5*b*. The cell thickness is about 0.13 mm. but is not exactly the same for each sample. It is seen that during the early stages of the reaction a new

absorption band appears at about 3300 cm.^{-1} , while the original amine band is reduced in intensity. This must mean that the presence of the anhydride is producing a shift in the frequency of the N-H band from 3320 to 3300 cm.^{-1} , and indicates that a proportion of the amine molecules are associated with the anhydride. A similar effect has been observed by Ashdown & Kletz (1948) in mixtures of alcohols and aldehydes. As the anhydride polymerizes the amount of associated amine decreases, until the absorption band assumes the contour of the original band obtained with the pure amine solution.

We wish to thank Mr A. H. Wilson for mathematical assistance, Dr C. H. Bamford and Dr M. J. S. Dewar for advice and encouragement, and Miss M. E. Henry for numerical solution of the differential equation (23).

REFERENCES

- Ashdown, A. & Kletz, T. A. 1948 *J. Chem. Soc.* p. 1454.
Astbury, W. T., Dalglish, C. E., Darmon, S. E. & Sutherland, G. B. B. M. 1948 *Nature*, **162**, 596.
Baughan, E. C. & Bell, R. P. 1937 *Proc. Roy. Soc. A*, **158**, 464.
Bell, R. P. & Trotman-Dickenson, A. F. 1949 *J. Chem. Soc.* p. 1288.
Elliott, A., Ambrose, E. J. & Temple, R. B. 1948 *J. Chem. Phys.* **16**, 877.
Flory, P. J. 1940 *J. Amer. Chem. Soc.* **62**, 1561.
Freudenberg, K. & Markert, L. 1927 *Ber. dtsch. chem. Ges.* **60 B**, 2447.
Houwink, R. 1940 *J. prakt. Chem.* **157**, 15.
Perry, S. & Hibbert, H. 1940 *J. Amer. Chem. Soc.* **62**, 2599.
Sigmund, F. & Wessely, F. 1926 *Hoppe-Seyl. Z.* **157**, 91.
Waley, S. G. 1948 *J. Chem. Soc.* p. 2008.

The use of quaternions in wave-tensor calculus

By C. W. KILMISTER, *Queen Mary College, London*

(Communicated by N. F. Mott, F.R.S.—Received 17 June 1949)

Quaternions are used to obtain some of Eddington's results in a general affine space, and a more convenient notation for *E*-numbers is suggested. As an example, Dirac's equations are derived, their connexion with the affine space of distant parallelism is explained, and Eddington's views on their 'Euclidean' nature are discussed.

1. INTRODUCTION

Eddington (1936, 1946) has given a far-reaching theory of physical phenomena in his unification of relativity theory and quantum theory; but the mathematical methods and explanations are not always conducive to ready understanding, and, in particular, the connexion with general relativity is not clear. Silberstein (1912, 1913), Lanczos (1929), Conway (1937), Fischer (1940) and others have shown the connexion of quaternions with special relativity, and Conway (1937) has suggested their use with Dirac's equation. Flint (1928), Podolsky (1931), Haskey (1940) and others have considered absolute parallelism in connexion with Dirac's equation.

In this paper an ennuple of vectors is used, in § 2, to connect quaternion algebra with a general affine space. In § 3 a linear quaternion function of a quaternion is defined, and alternative definitions are investigated. The equivalence of these functions to E -numbers is demonstrated in § 4, and a more convenient notation for E -numbers is suggested.

In § 5 the results already found are used to obtain Eddington's simple wave equation. In particular, the identical nature of this equation is shown to be obvious. Eddington's wave identities and the solution of the simple wave equation are found by direct methods.

In § 6, as an example, the theory is used to derive Dirac's equations by a modification of Eddington's method. The modification has the advantages:

- (i) of showing why distant parallelism enters into the consideration of these equations;
- (ii) of deriving the equations in the affine space of distant parallelism;
- (iii) of producing Dirac's equations in their usual form with four differential operators, instead of the sixteen required by Eddington (1936).

2. INTRODUCTION OF QUATERNIONS

In this section we adopt the convention that literal suffixes are to take the values 0 to 3, and we use the summation convention for them. In later sections these conventions apply only to Greek suffixes.

In a general affine space V_4 , there is set up at each point (x^0, x^1, x^2, x^3) , an ennuple of vectors $({}_i h^\alpha)$, which can also be described by the quantities $({}^i h_\alpha)$, where ${}_i h^\alpha$, ${}^i h_\alpha$ satisfy (Eisenhart 1927, p. 45)

$${}_i h^\alpha {}^i h_\beta = \delta^\alpha_\beta, \quad {}_i h^\alpha {}^j h_\alpha = \delta^j_i. \quad (2.1)$$

Here Latin suffixes denote vectors, Greek suffixes components relative to the co-ordinate system. The ennuple components of a tensor $T^{\alpha \dots \beta}_{\gamma \dots \delta}$ are the invariants $T^{i \dots j}_{k \dots l}$, where

$$T^{i \dots j}_{k \dots l} = {}^i h_\alpha \dots {}^j h_\beta {}_k h^\gamma \dots {}_l h^\delta T^{\alpha \dots \beta}_{\gamma \dots \delta}. \quad (2.2a)$$

We shall also require a mixed representation of a tensor in which certain of the suffixes refer to the co-ordinates, and others to the ennuple components. For example,

$$T^i_{k\alpha} = {}^i h_\beta {}_k h^\gamma T^\beta_{\gamma\alpha}. \quad (2.2b)$$

If T^μ is any contravariant vector, the equation

$$T_\mu = {}^i h_\mu {}^i h_\nu T^\nu \quad (2.3)$$

defines a covariant vector T_μ corresponding to T^μ . Operating on (2.3) with ${}_j h^\mu$ we have

$${}_j h^\mu T_\mu = {}^i h_\nu T^\nu,$$

and so by (2.2a),

$$T_j = T^j.$$

Hence corresponding vectors have equal ennuple components and so may be regarded as describing the same vector. Obviously ${}^i h_\alpha, {}^i h^\alpha$ are corresponding vectors. If we change to a new ennuple $({}_i h'^\alpha, {}^i h'_\alpha)$, where

$${}^i h'_\alpha = a_j^i {}^j h_\alpha, \quad {}^i h'^\alpha = A_i^j {}^j h^\alpha, \quad (2.4)$$

a_j^i, A_i^j being functions of position, then by (2.1), (2.4),

$$\begin{aligned} \delta_i^j &= {}^i h'^\alpha {}^j h'_\alpha = A_i^k a_l^j {}^k h^\alpha {}^l h_\alpha \\ &= A_i^k a_k^j. \end{aligned} \quad (2.5)$$

Hence A_i^j, a_j^i are reciprocal matrices; moreover, if we impose the condition that ${}_i h'^\alpha, {}^i h'_\alpha$ shall be a pair of corresponding vectors in the sense of (2.3), and we also use (2.1), we have

$$a_j^i {}^j h_\alpha = {}^k h_\alpha {}^k h_\beta A_i^j {}^j h^\beta = {}^i h_\alpha A_i^j.$$

Hence
$$a_j^i = A_i^j, \quad (2.6)$$

and equations (2.5), (2.6) imply that a_j^i is an orthogonal transformation of four variables.

Let $e_i (i = 0, 1, 2, 3)$ represent the vector ${}^i h_\alpha$ (or ${}_i h^\alpha$) and define multiplication by the equations

$$\left. \begin{aligned} e_0 e_i &= e_i e_0 = e_i \quad (i = 0, 1, 2, 3), \\ e_i e_j &= -e_j e_i = e_k \quad (\text{if } i, j, k \text{ is a cyclic permutation of } 1, 2, 3), \\ e_i e_j &= -e_0 \quad (\text{if } i = j \neq 0), \\ \lambda e_i &= e_i \lambda \quad (\text{if } \lambda \text{ is a scalar } i = 0, 1, 2, 3), \end{aligned} \right\} \quad (2.7)$$

the multiplication being distributive over addition. Since the matrices

$$e_0 = \begin{bmatrix} 1 & \cdot \\ \cdot & 1 \end{bmatrix}, \quad e_1 = \begin{bmatrix} i & \cdot \\ \cdot & -i \end{bmatrix}, \quad e_2 = \begin{bmatrix} \cdot & -1 \\ 1 & \cdot \end{bmatrix}, \quad e_3 = \begin{bmatrix} \cdot & -i \\ -i & \cdot \end{bmatrix}, \quad (2.8)$$

give a particular representation of the e_i , multiplication is associative and e_0 is the unit element. Except in § 6, where differential properties are involved, it is therefore permissible to write $e_0 = 1$. The equations (2.7) then define quaternion algebra, an account of which, suitable for our purposes, will be found in Brand (1947, chap. x). Brand, however, and most other authors, treat only of quaternions with components from the field of real numbers; those who consider quaternions with complex coefficients (e.g. Tait 1873, p. 64; Blaton 1935, p. 340) assume that the results established for real quaternions still hold, although null divisors now exist. In fact, the proofs of the equations found in §§ 181–183 of Brand still hold except that, for (183.1), the condition $N(q) \neq 0$ is needed.

It is convenient to modify Brand's notation for the conjugate, $\bar{K}q$, and the norm, Nq , of a quaternion, as follows:

$$\left. \begin{aligned} \bar{q} &= \text{conjugate of } q, \\ |q|^2 &= \text{norm of } q \quad (\text{Real } |q| \geq 0). \end{aligned} \right\} \quad (2.9)$$

Thus Brand (182·7) becomes

$$|pq| = \pm |p| \cdot |q|. \quad (2\cdot10)$$

There is now a (1:1) correspondence between corresponding vectors in V_4 and quaternions. For, since $T_i = T^i$, the vector described by either T_μ or T^μ satisfies

$$(T_\mu, T^\mu) = \sum_{i=0}^3 T_i({}^i h_\mu, {}^i h^\mu),$$

so that to the vector T_μ or T^μ corresponds the quaternion

$$T = \sum_{i=0}^3 T_i e_i. \quad (2\cdot11)$$

The ennuple transformation which leaves ${}^i h^\alpha$ and ${}^i h_\alpha$ corresponding vectors is orthogonal, and Lanczos (1929, p. 451) has shown that it may be written

$$q' = aqb \quad (|a| = |b| = 1). \quad (2\cdot12)$$

If we call any quaternion which transforms like q covariant, then any quaternion, say r , which transforms like $\bar{q}$ will be called contravariant. Quaternions which transform like $qr = s$, $rq = t$, respectively, will be called left-hand or right-hand mixed quaternions (Silberstein 1912). Hence

$$\left. \begin{aligned} q' &= aqb, \\ r' &= \bar{b}r\bar{a} \quad (\text{by Brand, 182}\cdot8), \\ s' &= as\bar{a} \quad (\text{by Brand, 182}\cdot3), \\ t' &= \bar{b}tb \quad (\text{by Brand, 182}\cdot3). \end{aligned} \right\} \quad (2\cdot13)$$

If in such a transformation there are quaternions ψ , ϕ such that

$$\psi' = a\psi, \quad \phi' = \phi b, \quad (2\cdot14)$$

ψ , ϕ will be called 'binions', of the first and second kind, respectively [German: 'Semiquaternionen', Blaton (1935)].

3. LINEAR FUNCTIONS

If q is any quaternion then a general ennuple transformation (2·4) is of the form

$$q' = f(q), \quad (3\cdot1)$$

where $f(q)$ is a function of q such that for any two quaternions q and r

$$f(q+r) = f(q) + f(r). \quad (3\cdot2)$$

Any f satisfying (3·1), (3·2) is called a 'linear function' (Hamilton 1866, p. 436; Joly 1903). If λ is a scalar, in the sense of quaternion algebra, it follows that

$$f(\lambda q) = \lambda f(q), \quad f(\lambda) = \lambda f(1), \quad f(0) = 0. \quad (3\cdot3)$$

The linear functions f and g are equal, provided that

$$f(q) = g(q) \quad (3\cdot4)$$

holds for all q . If f and g are any two linear functions, and

$$h(q) = f(q) + g(q),$$

then, by (3.2), $h(q)$ is also a linear function and we write

$$h = f + g. \quad (3.5)$$

If, also,

$$k(q) = f\{g(q)\},$$

then, by (3.2),

$$\begin{aligned} k(q+r) &= f\{g(q+r)\} \\ &= f\{g(q) + g(r)\} \\ &= f\{g(q)\} + f\{g(r)\} \\ &= k(q) + k(r). \end{aligned}$$

Hence, by (3.2), k is a linear function, which we write as

$$k = fg. \quad (3.6)$$

From these definitions, it follows that, for any three linear functions f , g , h , and a scalar λ ,

$$\left. \begin{aligned} f+g &= g+f, & (i), & & f+(g+h) &= (f+g)+h, & (ii) \\ (f+g)h &= fh+gh, & & & f(gh) &= (fg)h, & (iv) \\ f(g+h) &= fg+fh, & (iii), & & f\lambda &= \lambda f, & (v) \end{aligned} \right\} \quad (3.7)$$

As an example, we prove one of the equations (iii). If q is any quaternion, by (3.6), (3.5) and (3.2),

$$\begin{aligned} \{f(g+h)\}(q) &= f\{(g+h)(q)\} = f\{g(q) + h(q)\} \\ &= f\{g(q)\} + f\{h(q)\} \\ &= (fg)(q) + (fh)(q) \\ &= (fg+fh)(q). \end{aligned}$$

Hence, by (3.4),

$$f(g+h) = fg+fh.$$

We shall follow Hamilton (1866, p. 436) henceforward, and write fq instead of $f(q)$.

The following four definitions of a linear function are equivalent to (3.2):

Let $r = r_\alpha e_\alpha$, $q = q_\alpha e_\alpha$, the r_α , q_α being the scalar components of r , q . Then

$$\begin{aligned} (i) \quad & r = fq \quad \text{if} \quad r_\alpha = f_{\alpha\beta} q_\beta; \\ (ii) \quad & r = fq \quad \text{if} \quad \begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ r_3 \end{bmatrix} = \{f\} \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix}, \end{aligned}$$

where $\{f\}$ is a square matrix;

$$(iii) \quad fq = \sum_{i=0}^3 a_i S(a'_i q), \quad (3.8)$$

where a_i , a'_i are two sets of four quaternions, and S has the usual quaternion meaning

$$S(q_\alpha e_\alpha) = q_0.$$

This will be called 'Gibbs's form' from the analogy with Gibbs's dyadic (Gibbs 1884, chap. 3), but it is due to Hamilton (1866, p. 493).

$$(iv) \quad fq = \sum_{i=0}^3 a_i q a'_i, \quad (3.9)$$

where a_i, a'_i are two sets of four quaternions (not identical with those of definition (iii)). This will be called Hamilton's form; he assumed it as sufficiently general (Hamilton 1866, p. 488). These functions certainly satisfy (3.2); it remains to show that each form includes all linear functions. We have, by (3.2), (3.3),

$$fq = f(q_\alpha e_\alpha) = q_\alpha f e_\alpha,$$

$$\text{By (3.1) and Brand (181.6)} \quad f e_\alpha = f_{\beta\alpha} e_\beta,$$

where the $f_{\beta\alpha}$ are sixteen scalars. Hence

$$fq = (f_{\beta\alpha} q_\alpha) e_\beta$$

or

$$r_\alpha = f_{\alpha\beta} q_\beta,$$

giving forms (i) and (ii).

Again, define

$$\left. \begin{aligned} (i) &= e_i^2 \\ (i)(j) \dots (k) &= (ij \dots k); \end{aligned} \right\} \quad (3.10)$$

then we may write

$$q_i = (i) S(e_i q), \quad (3.11)$$

and so

$$\begin{aligned} fq &= f_{\beta\alpha} e_\beta q_\alpha = \sum_{\alpha=0}^3 (\alpha) f_{\beta\alpha} e_\beta S(e_\alpha q) \\ &= a_\alpha S(a'_\alpha q), \end{aligned}$$

which is form (iii) if $a_\alpha = (\alpha) f_{\beta\alpha} e_\beta$ (not summed for α),

$$a'_\alpha = e_\alpha.$$

To obtain form (iv), consider the linear function

$$\sigma q = q - e_1 q e_1 - e_2 q e_2 - e_3 q e_3 \quad (3.12)$$

We have

$$\sigma e_1 = \sigma e_2 = \sigma e_3 = 0, \quad \sigma e_0 = 4,$$

so that

$$\sigma q = 4Sq. \quad (3.13)$$

Hence

$$\begin{aligned} fq &= \sum_{\alpha=0}^3 (\alpha) f_{\beta\alpha} e_\beta S(e_\alpha q) \\ &= \frac{1}{4} \sum_{\alpha=0}^3 (\alpha) f_{\beta\alpha} e_\beta \sigma(e_\alpha q), \end{aligned}$$

which reduces, by (3.12), to

$$fq = \sum_{i,j=0}^3 f_{ij} e_i q e_j, \quad (3.14)$$

where the f_{ij} are sixteen coefficients, not equal to $f_{\alpha\beta}$. This is form (iv) with, for instance,

$$a_i = e_i, \quad a'_i = \sum_{j=0}^3 f_{ij} e_j.$$

We may write (3.14) as
$$f = \sum_{i,j=0}^3 f_{ij} K_{ij}, \quad (3.15)$$

where K_{ij} is the linear function defined by

$$K_{ij}q = e_i q e_j.$$

Unless otherwise stated, $f_{\alpha\beta}$ will mean the 'matrix elements' of $\{f\}$ in the form (ii) and f_{ij} the components of f in (3.14), (3.15).

4. CONNEXION WITH E -NUMBERS

The four anticommuting operators

$$iK_{11}, \quad iK_{21}, \quad iK_{31}, \quad K_{02} \quad (4.1)$$

have squares -1 and generate the whole set of K_{ij} . Hence (Eddington 1936, p. 21) the algebra of the K_{ij} , but for some factors $\pm i$, is isomorphic with the algebra of the E_μ (Conway 1937). Hence, by (3.15), the algebra of E -numbers is the algebra of linear functions of quaternions. This being so, the result mentioned by McCrea (1938), that every E -number satisfies a certain biquadratic, has already been proved without the aid of matrices by Hamilton (1866, p. 491).

The K_{ij} form a very convenient double-suffix notation for the E_μ . The multiplication rule is obviously

$$K_{ij}K_{kl} = K_{pq}, \quad (4.2)$$

where

$$e_i e_k = e_p, \quad e_i e_j = e_q,$$

and we also write

$$e_k e_i = -e_p \quad \text{as} \quad e_{-p},$$

so that

$$K_{i,-j} = K_{-i,j} = -K_{ij}.$$

For example, $K_{23}K_{12} = K_{-3,1} = -K_{31}, \quad K_{01}K_{23} = K_{22}.$

If $iK_{11}, iK_{21}, iK_{31}, K_{02} = E_1, E_2, E_3, E_4,$

then (Eddington 1936, equation 2.22)

$$K_{03} = E_5.$$

Hence

$$E_{15}, E_{25}, E_{35}, E_{45} = iK_{12}, iK_{22}, iK_{32}, -K_{01},$$

$$E_{14}, E_{24}, E_{34} = -iK_{13}, -iK_{23}, -iK_{33},$$

$$E_{12}, E_{23}, E_{32} = K_{30}, K_{10}, K_{20}.$$

Other identifications are, of course, possible. The matrix representation, S_{ij} , of K_{ij} may be found as in the following example. Let

$$i = 2, \quad j = 3, \quad r = K_{23}q,$$

so that

$$r = e_2 q e_3 = q_0 e_1 + q_1 - q_2 e_3 - q_3 e_2.$$

hence

$$\begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} . & 1 & . & . \\ 1 & . & . & . \\ . & . & . & -1 \\ . & . & -1 & . \end{bmatrix} \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix},$$

and therefore

$$S_{23} = \begin{bmatrix} . & 1 & . & . \\ 1 & . & . & . \\ . & . & . & -1 \\ . & . & -1 & . \end{bmatrix}.$$

The S_{ij} may be catalogued by using a notation in which the column occupied by +1 or -1 in each row is written down, a bar denoting that -1 occurs. Thus $S_{23} = (2\bar{1}4\bar{3})$ and we find

$$\begin{aligned} S_{00} &= (1234) = -iE_{16} & S_{01} &= (\bar{2}14\bar{3}) = E_{04}; \\ S_{02} &= (\bar{3}412) = E_{45}; & S_{03} &= (\bar{4}3\bar{2}1) = E_{05}; \\ S_{10} &= (\bar{2}1\bar{4}3) = -E_{12}; & S_{11} &= (\bar{1}234) = iE_{35}; \\ S_{12} &= (\bar{4}3\bar{2}1) = -iE_{03}; & S_{13} &= (\bar{3}41\bar{2}) = -iE_{34}; \\ S_{20} &= (\bar{3}41\bar{2}) = E_{31}; & S_{21} &= (\bar{4}32\bar{1}) = -iE_{25}; \\ S_{22} &= (\bar{1}2\bar{3}4) = iE_{02}; & S_{23} &= (214\bar{3}) = iE_{24}; \\ S_{30} &= (\bar{4}321) = E_{23}; & S_{31} &= (3\bar{4}1\bar{2}) = -iE_{15}; \\ S_{32} &= (\bar{2}14\bar{3}) = iE_{01}; & S_{33} &= (\bar{1}23\bar{4}) = iE_{14}. \end{aligned}$$

In this table the matrices are identified with the representation of the E -numbers given by Eddington (1946, p. 142).

In (3.15) the quarter-spur of f is defined by

$$qsf = f_{00}. \quad (4.3)$$

Hence

$$qs(f+g) = qsf + qsg, \quad qs(fg) = qs(gf), \quad (4.4)$$

and

$$f_{ij} = (ij) qs(K_{ij}f). \quad (4.5)$$

If $\{f\}$ is the matrix representation of f ,

$$\{f\} = \sum_{i,j=0}^3 f_{ij} S_{ij}.$$

But

$$\text{spur } S_{ij} = 0$$

unless $i = j = 0$, and

$$\text{spur } S_{00} = 4.$$

Hence

$$\text{spur } \{f\} = 4f_{00} = 4qsf. \quad (4.6)$$

We therefore obtain, using the definition (3.8) for f ,

$$qsf = \frac{1}{4} \sum_{i=0}^3 S(a_i a'_i), \quad (4.7)$$

and, using the definition (3.9) for f ,

$$qsf = \sum_{i=0}^3 (Sa_i)(Sa'_i). \quad (4.8)$$

Using (3.13) in (4.8) gives

$$\begin{aligned} 16qsf &= \sum_{i,j,k=0}^3 (j) e_j a_i e_j \cdot (k) e_k a'_i e_k \\ &= \sum_{j,k=0}^3 (jk) K_{jk} f K_{jk}, \end{aligned}$$

a result which Eddington (1936, equation 2.51) obtains by a different method. We shall now prove a result related to Eddington's (1936, equation 3.37), that is, if

$$fq = \sum_{i=0}^3 a_i S(a'_i q),$$

then

$$f_{ij} = \sum_{k=0}^3 \frac{1}{4}(ij) S(e_i a_k e_j a'_k). \quad (4.10)$$

For, by (3.6),

$$K_{ij}fq = \sum_{k=0}^3 e_i a_k e_j S(a'_k q),$$

and so, by (4.7),

$$qs(K_{ij}f) = \sum_{k=0}^3 \frac{1}{4} S(e_i a_k e_j a'_k).$$

Hence, using (4.5), the result (4.10) follows. If f is 'pure' (Eddington 1936, p. 65), then in our notation we have (say)

$$fq = aS(a'q),$$

and so, by (4.10),

$$f_{ij} = \frac{1}{4}(ij) S(e_i a e_j a'). \quad (4.11)$$

5. THE SIMPLE WAVE EQUATION

Let ψ, χ be two quaternions; then, by Brand (181.10),

$$\psi\chi \cdot \psi = \psi \cdot \chi\psi.$$

But, by Brand (181.13), $S(\psi\chi) \cdot \psi = \psi \cdot S(\chi\psi).$

Subtracting the second equation from the first,

$$p\psi = \psi q, \quad (5.1)$$

where

$$p = V \cdot \psi\chi, \quad q = V \cdot \chi\psi,$$

with the usual quaternion notation,

$$Vr = r - Sr.$$

Interchanging ψ, χ , we have

$$q\chi = \chi p. \quad (5.2)$$

If we write in (5.1)

$$p = \sum_{i=1}^3 \varpi_{i1} e_i, \quad q = -me_1 + \varpi_{03} e_2 - \varpi_{02} e_3,$$

and operate with K_{01} , we obtain

$$(\varpi_{11} K_{11} + \varpi_{21} K_{21} + \varpi_{31} K_{31} + \varpi_{02} K_{02} + \varpi_{03} K_{03} - m) \psi = 0.$$

From (4.1), this is Eddington's simple wave equation (1936, §5.4) which, when written in the form (5.1), is obviously an identity.

Conversely, if ψ, χ satisfy (5.1) and (5.2), and p and q are vectors, then

$$p^2\psi = p\psi q = \psi q^2.$$

But p^2, q^2 are scalars; hence, if $\psi \neq 0$,

$$p^2 = q^2. \quad (5.3)$$

Also

$$p\psi\chi = \psi q\chi = \psi\chi p;$$

hence (Brand (181.12)), if $V.\psi\chi \neq 0$, we have

$$p = \lambda V.\psi\chi, \quad q = \lambda V.\chi\psi. \quad (5.4)$$

The equation for q follows by interchanging ψ and χ , the constant λ being the same for each case, by (5.1).

If

$$p = \sum_{i=1}^3 p_i e_i, \quad q = \sum_{i=1}^3 q_i e_i,$$

then using (3.11), (5.4) and the fact that $S(e_i S.\psi\chi) = 0$, we find

$$\begin{aligned} p_i &= -S(e_i p) = -\lambda S(e_i V.\psi\chi) \\ &= -\lambda S(e_i \psi\chi). \end{aligned}$$

If

$$j_r = \psi S(\chi r),$$

then, by (4.11),

$$\left. \begin{aligned} p_i &= 4\lambda j_{i0}, \\ q_i &= 4\lambda j_{0i}. \end{aligned} \right\} \quad (5.5)$$

and similarly

$$\text{Hence} \quad \frac{j_{10}}{p_1} = \frac{j_{20}}{p_2} = \frac{j_{30}}{p_3} = \frac{j_{01}}{q_1} = \frac{j_{02}}{q_2} = \frac{j_{03}}{q_3} = \frac{1}{4\lambda}, \quad (5.6)$$

and, by (5.3) and (5.5),

$$j_{10}^2 + j_{20}^2 + j_{30}^2 = j_{01}^2 + j_{02}^2 + j_{03}^2,$$

results equivalent to Eddington's (1936, equations (5.33), (5.45)).

Define now, for $i = 0, 1, 2, 3$,

$$p^{(i)} = \lambda V(\psi e_i \chi), \quad q^{(i)} = \lambda V(\chi e_i \psi), \quad (5.7)$$

so that

$$p^{(0)} = p, \quad q^{(0)} = q.$$

Then, as in (5.5),

$$\left. \begin{aligned} p_j^{(i)} &= 4\lambda(i)j_{ji} \\ q_j^{(i)} &= 4\lambda(i)j_{ij} \end{aligned} \right\} \quad (j = 1, 2, 3). \quad (5.8)$$

But, by (5.1),

$$S(\chi p \psi e_i) = S(\chi \psi q e_i).$$

Using Brand (181.9) for $\chi\psi$, and since p is a vector,

$$S\{p.V(\psi e_i \chi)\} = S(\chi\psi) \cdot S(qe_i) + S\{(V.\chi\psi)qe_i\},$$

and therefore by (5.5), (5.8), (4.10),

$$S.p p^{(i)} = 4\lambda j_{00}(i)q_i + S(q^2 e_i).$$

If $i = 0$,

$$S p^2 = S q^2,$$

which is (5.3) again; if $i \neq 0$,

$$S.p p^{(i)} = -4\lambda j_{00} q_i,$$

and so

$$\sum_{j=1}^3 p_j p_j^{(i)} - 16\lambda^2 j_{00} j_{0i} = 0,$$

from which it follows that

$$j_{10}j_{1i} + j_{20}j_{2i} + j_{30}j_{3i} + j_{00}j_{0i} = 0, \quad (5.9)$$

corresponding to Eddington's (1936, equation 5.46).

We shall now show that the simple wave equation, in the form (5.1), has a solution

$$\psi = p\gamma + \gamma q, \quad (5.10)$$

where γ is an arbitrary vector. For, by (5.3),

$$\begin{aligned} p\psi - \psi q &= p(p\gamma + \gamma q) - (p\gamma + \gamma q)q \\ &= (p^2 - q^2)\gamma = 0. \end{aligned}$$

Tait (1873, p. 97) has shown that this is the most general solution. The solution for χ is similarly

$$\chi = q\delta + \delta p, \quad (5.11)$$

where δ is another arbitrary vector (cf. Eddington 1936, equation 5.26).

$$\text{Writing (5.1) as} \quad H\psi = 0, \quad (5.12)$$

$$\text{let} \quad \psi' = f\psi,$$

where f is a linear function. Equation (5.12) will be invariant under the transformation from ψ to ψ' provided that

$$H' = hHf^{-1}, \quad (5.13)$$

where h is a linear function having a reciprocal. In particular, let ψ be a binion of the first kind; then, by (2.14),

$$\begin{aligned} H'\psi &= h(pa^{-1}\psi - a^{-1}\psi q) \\ &= (ha^{-1})(apa^{-1}\psi - \psi q). \end{aligned} \quad (5.14)$$

Hence p, q behave like left-hand mixed, and invariant, quaternions, respectively, independently of h .

6. DIFFERENTIAL OPERATIONS

As an example of the foregoing theory, it will be applied to the establishment of Dirac's equations, following Eddington's method (1936, chap. 8).

Assumption 1. We assume the space V_4 is affinely-connected; if the affine connexion has coefficients $L_{\nu\sigma}^\mu$, the covariant derivatives of the ennuple vectors are given by (Eisenhart 1927 (3.2))

$$\left. \begin{aligned} {}_i h^\alpha_{,\beta} &= \frac{\partial}{\partial x^\beta} ({}_i h^\alpha) + L_{\nu\beta}^\alpha {}_i h^\nu, \\ {}_i h_{\alpha,\beta} &= \frac{\partial}{\partial x^\beta} ({}_i h_\alpha) - L_{\alpha\beta}^\nu {}_i h_\nu. \end{aligned} \right\} \quad (6.1)$$

$$\begin{aligned} \text{Define} \quad \Delta_{\beta\gamma}^\alpha &= \sum_{i=0}^3 {}_i h^\alpha \frac{\partial}{\partial x^\gamma} ({}_i h_\beta) \\ &= - \sum_{i=0}^3 {}_i h_\beta \frac{\partial}{\partial x^\gamma} ({}_i h^\alpha). \end{aligned} \quad (6.2)$$

by (2.1). We now introduce quantities $\Lambda_{\nu\beta}^\alpha$, where

$$\left. \begin{aligned} {}_i h^\alpha_{,\beta} &= (L_{\nu\beta}^\alpha - \Delta_{\nu\beta}^\alpha) {}_i h^\nu = \Lambda_{\nu\beta}^\alpha {}_i h^\nu, \\ {}_i h_{\alpha,\beta} &= -\Lambda_{\alpha\beta}^\nu {}_i h_\nu. \end{aligned} \right\} \quad (6.3)$$

and

Hence the $\Lambda_{\nu\beta}^\alpha$ are components of a tensor which, by (2.2b), can also be written in a mixed form. It follows that

$$\left. \begin{aligned} {}_i h^\alpha_{,\beta} &= \sum_{j=0}^3 \Lambda_{i\beta}^j {}_j h^\alpha, \\ {}^i h_{\alpha,\beta} &= - \sum_{j=0}^3 \Lambda_{j\beta}^i {}^i h_\alpha. \end{aligned} \right\} \quad (6.4)$$

Assumption 2. In order to limit our discussion to the field of quaternions, we assume the covariant derivative of e_i is a quaternion,

$$e_{i,\beta} = \sum_{k=0}^3 \Pi_{i\beta}^k e_k.$$

Since the ${}_i h^\alpha$ (or ${}^i h_\alpha$) and e_i represent the same vector, we have, by (6.4),

$$\Pi_{i\beta}^k = \Lambda_{i\beta}^k = -\Lambda_{k\beta}^i,$$

and therefore

$$\Lambda_{i\beta}^j + \Lambda_{j\beta}^i = 0, \quad (6.5)$$

while, by (2.2),

$$e_{i,j} = \sum_{k=0}^3 \Lambda_{ij}^k e_k. \quad (6.6)$$

Assumption 3. We assume, further, that

$$(qr)_{,i} = q_{,i} r + q r_{,i} \quad (6.7)$$

for any two quaternions q, r ; this imposes a further restriction on the Λ_{ij}^k . In particular,

$$e_0 e_i = e_i,$$

so that

$$e_{0,j} e_i + e_0 e_{i,j} = e_{i,j} = e_0 e_{i,j},$$

and hence, for all j ,

$$e_{0,j} = 0.$$

Thus from (6.6) we have

$$\Lambda_{0j}^k = \Lambda_{kj}^0 = 0. \quad (6.8)$$

Now using (6.5) and (6.8) we obtain

$$\begin{aligned} e_{1,j} &= \Lambda_{1j}^2 e_2 + \Lambda_{1j}^3 e_3 \\ &= \beta_j e_1 - e_1 \beta_j, \end{aligned} \quad (6.9)$$

where

$$2\beta_j = \Lambda_{2j}^3 e_1 + \Lambda_{3j}^1 e_2 + \Lambda_{1j}^2 e_3 \quad (j = 0 \dots 3). \quad (6.10)$$

By symmetry,

$$e_{i,j} = \beta_j e_i - e_i \beta_j \quad (i, j = 0 \dots 3). \quad (6.11)$$

Hence, if

$$q = q_\alpha e_\alpha, \quad r = r_\alpha e_\alpha,$$

we have, using (6.11) and the fact that the covariant derivative of a scalar is its ordinary derivative,

$$\begin{aligned} (qr)_{,i} &= (q_\alpha r_\beta e_\alpha e_\beta)_{,i} \\ &= \frac{\partial}{\partial x^i} (q_\alpha r_\beta) \cdot e_\alpha e_\beta + q_\alpha r_\beta (\beta_i e_\alpha e_\beta - e_\alpha e_\beta \beta_i) \\ &= \frac{\partial q_\alpha}{\partial x^i} r_\beta \cdot e_\alpha e_\beta + q_\alpha \frac{\partial r_\beta}{\partial x^i} e_\alpha e_\beta + q_\alpha \cdot r_\beta \{ (\beta_i e_\alpha - e_\alpha \beta_i) e_\beta + e_\alpha (\beta_i e_\beta - e_\beta \beta_i) \} \\ &= q_{,i} r + q r_{,i}. \end{aligned}$$

Hence (6.5) and (6.8) are necessary and sufficient conditions for (6.7). From (6.8) we have

$${}_0h^\alpha_{,\beta} = 0. \quad (6.12)$$

In order that (6.12) should be valid for ennuple transformations which alter ${}_0h^\alpha$, there must exist a field of contravariant vectors parallel to any given vector. Hence (Eisenhart 1927, p. 19) the curvature tensor for the connexion (6.1) is zero, and therefore if the connexion is asymmetric, we have (Eisenhart 1927, p. 48),

$$L^\mu_{\nu\sigma} = \sum_{i=0}^3 {}^i\lambda^\mu \frac{\partial}{\partial x^\sigma} ({}^i\lambda_\nu), \quad (6.13)$$

where the ${}^i\lambda^\mu$ (or ${}^i\lambda_\mu$) are components of an ennuple of vectors. This is the affine space of distant parallelism, the vectors ${}^i\lambda^\mu$, ${}^i\lambda_\mu$ determining the parallelism. If now these vectors are expressed in terms of the ennuple ${}_ih^\alpha$, ${}_ih_\alpha$, we have

$${}^i\lambda_\alpha = \sum_{j=0}^3 a^i_j {}^j h_\alpha, \quad {}^i\lambda^\alpha = \sum_{j=0}^3 A^i_{ij} {}^j h^\alpha,$$

where, as in (2.5), $\sum_{k=0}^3 A^k_i a^i_k = \delta^i_i$. We find, after some calculation, using (6.13), (6.3) and (6.2),

$$\Lambda^i_{jk} = \sum_{p=0}^3 A^i_p {}^k h^\sigma \frac{\partial}{\partial x^\sigma} (a^p_j), \quad (6.14)$$

and hence

$$\frac{\partial}{\partial x^\sigma} (a^p_j) = \sum_{i,k=0}^3 {}^k h_\sigma a^p_i \Lambda^i_{jk},$$

If the arbitrary h -ennuple is chosen as the ennuple describing the parallelism, so that

$$\Lambda^i_{jk} = 0, \quad (6.15)$$

then (6.14) shows that only ennuple transformations with constant coefficients will leave the parallelism unchanged. Such transformations will be called 'permissible'. In this space of distant parallelism, if q is a quaternion,

$$q = \sum_{i=0}^3 q^i e_i,$$

we have, by (2.2a),

$$\begin{aligned} q^i_{,k} &= {}^k h^\sigma ({}^i h_\mu q^\mu)_\sigma \\ &= {}^k h^\sigma \frac{\partial}{\partial x^\sigma} ({}^i h_\mu q^\mu) \\ &= \frac{\partial q^i}{\partial X^k}, \end{aligned} \quad (6.16)$$

where, since dx^μ is a vector, it has ennuple components which we shall write as

$$dX^i = {}^i h_\sigma dx^\sigma,$$

defining a corresponding differential operator by

$$\frac{\partial}{\partial X^i} = {}^i h^\sigma \frac{\partial}{\partial x^\sigma}. \quad (6.17)$$

Hence, by (6.15),
$$q_{,k} = \frac{\delta q}{\delta X^k}, \quad (6.18)$$

where $\delta/\delta X^k$ denotes differentiation treating the e_i as constants. Hence, under permissible transformations, $q_{,k}$ transforms like q .

Now define
$$Dq = \sum_{i=0}^3 (i) e_i q_{,i},$$

so that
$$S \cdot Dq = \frac{\partial q_0}{\partial X^0} + \frac{\partial q_1}{\partial X^1} + \frac{\partial q_2}{\partial X^2} + \frac{\partial q_3}{\partial X^3}.$$

If Dq is a mixed quaternion (e.g. if q and e_i are mixed quaternions of the same kind), $S \cdot Dq$ is invariant. In such a case, we may define

$$\operatorname{div} q = S \cdot Dq,$$

so that
$$\operatorname{div} q = S \left(\frac{\delta q}{\delta X^0} - \nabla q \right) \quad (6.19)$$

where
$$\nabla q = \sum_{i=1}^3 e_i \frac{\delta q}{\delta X^i}.$$

It is evident that
$$S \left\{ \frac{\delta}{\delta X^0} (\nabla q) \right\} = S \{ \nabla (Sq) \} = 0. \quad (6.20)$$

Assumption 4. In § 5 we assume ψ is a binion of the first kind. It follows from (5.5) and (5.14) that $\sum_{i=1}^3 j_{i0} e_i$ is a left-hand mixed quaternion, and $j_{0i} (i = 1, 2, 3)$ are three invariants. Hence

$$q(\beta) = j_{0\beta} + \sum_{i=1}^3 j_{i0} e_i \quad (\beta = 1, 2, 3)$$

are three left-hand mixed quaternions. If one of these is identified with the four-dimensional probability vector (cf. Eddington 1936, p. 118), then, for $\beta = 1, 2$ or 3 ,

$$\operatorname{div} q(\beta) = 0.$$

Using (4.10), (5.4), (6.19) and (6.20) this becomes

$$S \left\{ \frac{\delta}{\delta X^0} (\psi f_\beta \chi) - \nabla (\psi \chi) \right\} = 0, \quad (6.21)$$

where f_β is one of e_1, e_2 or e_3 , but is written with a different letter to emphasize that it is invariant under the transformation.

Now
$$\nabla (\psi \chi) = \sum e_i (\psi_{,i} \chi + \psi \chi_{,i}),$$

so that
$$S \{ \nabla (\psi \chi) \} = S \{ (\nabla \psi) \chi + \psi (\chi \nabla') \},$$

where
$$\chi \nabla' = \sum_{i=1}^3 \chi_{,i} e_i = \sum_{i=1}^3 \frac{\delta \chi}{\delta X^i} e_i.$$

Hence
$$S \left\{ \left(\frac{\delta \psi}{\delta X^0} f_\beta - \nabla \psi \right) \chi + \psi \left(f_\beta \frac{\delta \chi}{\delta X^0} - \chi \nabla' \right) \right\} = 0.$$

But $\nabla\psi - \frac{\delta\psi}{\delta X^0}f_\beta$ is some linear function of ψ , so that we may write

$$\left. \begin{aligned} \nabla\psi - \frac{\delta\psi}{\delta X^0}f_\beta &= g_{(\beta)}\psi, \\ \chi\nabla' - f_\beta \frac{\delta\chi}{\delta X^0} &= -h_{(\beta)}\chi, \end{aligned} \right\} \quad (6.22)$$

and similarly]

where $g_{(\beta)}$, $h_{(\beta)}$ are linear functions, which vary from point to point. Hence, finally,

$$S\{\chi \cdot g_{(\beta)}\psi - \psi \cdot h_{(\beta)}\chi\} = 0.$$

One way of satisfying this equation (cf. Eddington 1936, p. 119) is to make $h_{(\beta)}$ the conjugate function to $g_{(\beta)}$, that is, $\bar{g}_{(\beta)}$ (Hamilton 1866, p. 436). Hence

$$\left. \begin{aligned} \nabla\psi - \frac{\delta\psi}{\delta X^0}e_\beta &= g_{(\beta)}\psi, \\ \chi\nabla' - e_\beta \frac{\delta\chi}{\delta X^0} &= -\bar{g}_{(\beta)}\chi, \end{aligned} \right\} \quad (6.23)$$

where we now replace f_β by e_β since there is no longer any question of transformation. With a suitable choice of the function $g_{(\beta)}$ (Eddington 1936, equation 8.261), these are Dirac's equations in the form assumed by Conway (1937, equation 3.7). Using (6.17) the first equation (for example) may be written

$$\sum_{j=0}^3 \left[\left(\sum_{i=1}^3 e_i e_j h^\sigma - e_j e_\beta h^\sigma \right) \frac{\partial \psi_j}{\partial x^\sigma} \right] = \gamma_{\mu\nu} e_\mu \psi_\nu,$$

where $\gamma_{\mu\nu}$ are the matrix elements of $g_{(\beta)}$, and $\psi = \sum_{j=0}^3 \psi_j e_j$. With the notation

$$p_i = i h^\sigma \frac{\partial}{\partial x^\sigma}$$

these are seen to assume the usual form for the equations, written at length.

This derivation of Dirac's equations rests on four assumptions, of which the first limits consideration to affine spaces and the fourth specifies the kind of variables which are to appear as wave functions. The second and third are necessary conditions for the covariant differentiation of quaternions; they may be replaced by the assumption of an affine space of distant parallelism. Equation (6.23) shows that Dirac's equations in this space of distant parallelism are identical in form with the usual equations in Euclidean space. The fact that the equations are independent of any metric seems to be the correct interpretation of Eddington's statement that 'attempts to "extend Dirac's equation to general relativity" are misguided' (Eddington 1946, p. 12, footnote). Similarly, the fact that 'We deal only with a flat space in which a normal metric is always available' (Eddington 1932) has not affected the generality of Eddington's results on the differential wave equation. It seems possible that Eddington may have confused the need for a space in which the curvature tensor with respect to the affine connexion is zero (that is, an affine space of distant parallelism), with the need for a *flat* space, i.e. one in which the Riemann-Christoffel tensor is zero. It would be surprising if it were necessary for the space to

be flat, in order to establish equations which, as we have shown, have no connexion with metric.

The theory given in this section may be generalized to include non-permissible transformations which preserve (6.5) and (6.8); briefly stated the resultant equations are of the form (6.22), but the relation between $g_{(\beta)}$ and $h_{(\beta)}$ is more complicated.

The author wishes to thank Professor G. C. McVittie for suggesting this subject to him, and for helpful advice and encouragement in the course of the work.

REFERENCES

- Blaton, J. 1935 *Z. Phys.* **95**, 337.
 Brand, L. 1947 *Vector and Tensor Analysis*. New York: John Wiley.
 Conway, A. W. 1937 *Proc. Roy. Soc. A*, **162**, 145.
 Eddington, Sir A. 1932 *Proc. Roy. Soc. A*, **138**, 21.
 Eddington, Sir A. 1936 *Relativity theory of protons and electrons*. Cambridge University Press.
 Eddington, Sir A. 1946 *Fundamental theory*. Cambridge University Press.
 Eisenhart, L. P. 1927 *Non-Riemannian geometry*. New York: American Math. Society.
 Fischer, O. F. 1940 *Phil. Mag.* (7), **30**, 135.
 Flint, H. T. 1928 *Proc. Roy. Soc. A*, **121**, 676.
 Gibbs, W. 1884 *Vector analysis*. New Haven: Privately printed.
 (or *Collected works*. Longmans 1931).
 Hamilton, Sir W. R. 1866 *Elements of quaternions*. London: Longmans.
 Haskey, H. W. 1940 *Phil. Mag.* (7), **30**, 478.
 Joly, C. J. 1903 *Phil. Trans. A*, **201**, 223.
 Lanczos, von C. 1929 *Z. Phys.* **57**, 447.
 McCrea, W. H. 1938 *J. Lond. Math. Soc.* **13**, 283.
 Podolsky, B. 1931 *Phys. Rev.* **37**, 1398.
 Silberstein, L. 1912 *Phil. Mag.* (6), **23**, 790.
 Silberstein, L. 1913 *Phil. Mag.* (6), **25**, 135.
 Tait, P. G. 1873 *Quaternions*. Oxford: Clarendon Press.

Steady flow in the laminar boundary layer of a gas

By C. R. ILLINGWORTH, *Department of Mathematics, University of Manchester*

(Communicated by S. Goldstein, F.R.S.—Received 21 June 1949)

If the boundary-layer equations for a gas are transformed by Mises's transformation, as was done by Kármán & Tsien for the flow along a flat plate of a gas with unit Prandtl number σ , the computation of solutions is simplified, and use may be made of previously computed solutions for an incompressible fluid. For any value of the Prandtl number, and any variation of the viscosity μ with the temperature T , after the method has been applied to flow along a flat plate (a problem otherwise treated by Crocco), the flow near the forward stagnation point of a cylinder is calculated with dissipation neglected, both with the effect of gravity on the flow neglected and with this effect retained for vertical flow past a horizontal cylinder. The approximations involved by the neglect of gravity are considered generally, and the cross-drift is calculated when a horizontal stream flows past a vertical surface. When $\sigma = 1$, $\mu \propto T$, and the boundary is heat-insulated, it is shown that the boundary-layer equations for a gas may be made identical, whatever be the main stream, with the boundary-layer equations for an incompressible fluid with a certain, determinable, main stream. The method is also applied to free convection at a flat plate (with the heat of dissipation and the variation with altitude of the state of the surrounding fluid neglected) and to laminar flow in plane wakes, but for plane jets the conditions $\sigma = 1$, $\mu \propto T$, previously imposed by Howarth, are also imposed here in order to obtain simple solutions.

1. INTRODUCTION AND SUMMARY

Kármán & Tsien (1938), Crocco (1946), and Howarth (1948) have shown how, in some cases, steady flow in the laminar boundary layer of a compressible fluid can be easily calculated from a knowledge of the corresponding flow of an incompressible fluid. Kármán & Tsien use Mises's transformation of the boundary-layer equations to discuss the flow in the boundary layer at a flat plate in a uniform stream of gas parallel to the plate, with the Prandtl number $\sigma = 1$. They use the relation $\mu_r \propto T_r^{0.76}$ to express the viscosity μ_r as a function of the temperature T_r . Crocco, using as independent variables the distance x measured along the boundary and the velocity component parallel to the boundary, makes a thorough investigation of the boundary layer at a flat plate in a uniform stream for a range of values of σ . He considers both Sutherland's law for the viscosity as a function of the temperature and convenient approximations to it of the form $\mu_r \propto T_r^w$. In particular, when $\mu_r \propto T_r$, Crocco's equations reduce to those for an incompressible fluid. Howarth, assuming that $\sigma = 1$, $\mu_r \propto T_r$ in the flow past a heat-insulated boundary, uses as independent variables x and a distortion of the distance y normal to the boundary. His work is concerned chiefly with the boundary layer of a heat-insulated wall in a non-uniform stream, but he also considers the applications of his method to the flows in two-dimensional wakes and jets.

In this paper it is shown how, when the boundary-layer equations for a compressible fluid are transformed by Mises's transformation, the computation of solutions is simplified in several cases (including, of course, Kármán & Tsien's case of a uniform stream), and in particular how use may be made of previously computed solutions for an incompressible fluid. In some cases when $\mu_r \propto T_r$, the

equations for a compressible and for an incompressible fluid may be made identical (as in Crocco's work); in others the solution for a compressible fluid may be found by Kármán & Tsien's method of successive approximation, starting from a known solution for an incompressible fluid.

In § 3 Kármán & Tsien's method is extended to flow in the boundary layer at a flat plate when $\sigma \neq 1$, the viscosity being any given function of the temperature—the problem treated by Crocco (1946).

In § 4 the flow in the boundary layer is considered near the foremost generator of a blunt-nosed cylinder. Neither of the restrictions $\sigma = 1$, $\mu_r \propto T_r$ used by Howarth (1948) in his solution of this problem for a heat-insulated boundary needs to be imposed, but the dissipation is neglected, so the present solution is not applicable with a heat-insulated boundary. A special case is computed with $\sigma = 0.725$ and Sutherland's law for the variation of viscosity with temperature.

In § 5 the restrictions used by Howarth, that $\sigma = 1$, $\mu_r \propto T_r$, and that the boundary is heat-insulated, are imposed, and it is then shown that the boundary-layer equations for a compressible fluid may be made *identical, whatever be the main stream*, with the boundary-layer equations for an incompressible fluid with a certain, determinable, main stream.†

In §§ 3, 4 and 5 the effect of gravity on the flow is completely neglected. The approximations involved by the neglect of gravity, for flow past the surface of a cylinder, are discussed in §§ 2.31, 2.32, and the cross-drift when a horizontal stream flows past a vertical surface is computed in § 6 for a flat plate and for a blunt-nosed cylinder near its foremost generator. In § 7 it is shown that gravity may fairly easily be taken into account (though a special computation is necessary for each shape) in computing the flow near the foremost generator of a horizontal blunt-nosed cylinder in a vertical stream.

In § 8 the equations of free convection at a flat plate are reduced to an easily soluble form by the general methods used here, the heat of dissipation and the variation with altitude of the state of the surrounding fluid being neglected. This problem has previously been considered by Schmidt (1934) on the basis of equations that do not take full account of the compressibility of the fluid.

Finally, in § 9 simple solutions are found for laminar flow in plane jets and wakes. It is shown that, whereas for jets simple solutions cannot be easily obtained without the restrictions $\sigma = 1$, $\mu_r \propto T_r$, imposed by Howarth in his solutions for these flows, for wakes this is not the case.

2. THE BOUNDARY-LAYER EQUATIONS

2.1. *Formulation of the general equations*

Consider the boundary layer produced by the steady flow of a perfect gas past a cylindrical obstacle whose generators are perpendicular to the incident stream. Let the z -axis coincide with the foremost generator; let x be distance measured round a normal section of the cylinder, in the plane $z = \text{const.}$, from the foremost point, and let y be distance from the cylinder in the plane $z = \text{const.}$

† I am informed by Dr Howarth that Mr Stewartson has also obtained this result.

Values in the boundary layer being characterized by the suffix r , the equation of continuity becomes

$$\frac{\partial}{\partial x}(\rho_r u_r) + \frac{\partial}{\partial y}(\rho_r v_r) + \frac{\partial}{\partial z}(\rho_r w_r) = 0, \quad (1)$$

when the usual approximations of boundary-layer theory are made. The equations of momentum, when there are no body forces other than gravity, become

$$\rho_r D u_r = -g \rho_r l - \frac{\partial p_r}{\partial x} + \frac{\partial}{\partial y} \left(\mu_r \frac{\partial u_r}{\partial y} \right), \quad (2)$$

$$\frac{\partial p_r}{\partial y} = O(1), \quad (3)$$

$$\rho_r D w_r = -g \rho_r n - \frac{\partial p_r}{\partial z} + \frac{\partial}{\partial y} \left(\mu_r \frac{\partial w_r}{\partial y} \right), \quad (4)$$

where

$$D = u_r \frac{\partial}{\partial x} + v_r \frac{\partial}{\partial y} + w_r \frac{\partial}{\partial z},$$

and l, m, n are the direction cosines of the upward vertical relative to the orthogonal triad formed by the principal tangent, the normal, and the generator at the point $x, 0, z$. The equation of energy becomes

$$\rho_r D i_r - D p_r = \frac{1}{\sigma} \frac{\partial}{\partial y} \left(\mu_r \frac{\partial i_r}{\partial y} \right) + \mu_r \left[\left(\frac{\partial u_r}{\partial y} \right)^2 + \left(\frac{\partial w_r}{\partial y} \right)^2 \right], \quad (5)$$

and the equation of state is

$$p_r = \rho_r (i_r - e_r) = \bar{R} \rho_r T_r, \quad (6)$$

where

$$i_r = \int_0^{T_r} c_p(\tau) d\tau \quad \text{and} \quad e_r = \int_0^{T_r} c_v(\tau) d\tau$$

are the enthalpy and intrinsic energy of the gas. The specific heats, c_p and c_v , are functions of T_r only and $c_p - c_v$ is the gas constant $\bar{R}$. In equation (5) the Prandtl number σ has been assumed independent of the state of the gas. The variation of c_p and c_v with temperature is much less important than that of μ_r , which is also a function of T_r only, and in some cases it will be assumed that c_p and c_v are constant.

At the outside edge of the boundary layer, where values are characterized by the suffix 1, the velocity at the section x has the single component $u_1(x)$ in the direction of the tangent to the cylinder at $x, 0, z$. The quantities p_1, ρ_1 , etc., are functions of x and z only. The main stream is described by the equations

$$-\rho_1 u_1 \frac{du_1}{dx} = g \rho_1 l + \frac{\partial p_1}{\partial x}, \quad (7)$$

$$0 = g \rho_1 n + \frac{\partial p_1}{\partial z}, \quad (8)$$

corresponding to (2) and (4), and

$$\rho_1 \frac{\partial i_1}{\partial x} = \frac{\partial p_1}{\partial x}, \quad (9)$$

corresponding to (5), together with the equation of state

$$p_1 = \rho_1 (i_1 - e_1) = \bar{R} \rho_1 T_1. \quad (10)$$

From equation (3) we deduce in the usual way that we may take $p_r = p_1$, and it follows, by adding (7) to (2) and (8) to (4), that the momentum equations may be written

$$\rho_r D u_r - \rho_1 u_1 \frac{d u_1}{d x} = g(\rho_1 - \rho_r) l + \frac{\partial}{\partial y} \left(\mu_r \frac{\partial u_r}{\partial y} \right), \quad (11)$$

$$\rho_r D w_r = g(\rho_1 - \rho_r) n + \frac{\partial}{\partial y} \left(\mu_r \frac{\partial w_r}{\partial y} \right). \quad (12)$$

The energy equation (5) may be written, with the help of (9),

$$\rho_r D i_r - u_r \rho_1 \frac{\partial i_1}{\partial x} - w_r \frac{\partial p_1}{\partial z} = \frac{1}{\sigma} \frac{\partial}{\partial y} \left(\mu_r \frac{\partial i_r}{\partial y} \right) + \mu_r \left[\left(\frac{\partial u_r}{\partial y} \right)^2 + \left(\frac{\partial w_r}{\partial y} \right)^2 \right]. \quad (13)$$

2.2. The influence of gravity

The equations (1) and (7) to (13) are general, but in order to solve them conditions will henceforward be assumed independent of z . Before proceeding to solutions, however, the implications of this assumption will be considered. The flow is certainly assumed to be uninfluenced by the ends of the cylinder. There are, however, also conditions to be considered involving the effect of gravity on the flow.

Gravity causes a variation of the state of the main stream with altitude, and in particular, unless $n = 0$, a variation of p_1 , ρ_1 , etc., with z . In the exceptional case, $n = 0$, the generators of the cylinder are horizontal. Unless $n = 0$, gravity also causes a transverse flow in the boundary layer with velocity w_r parallel to the generators; the pressure across a section of the boundary layer is approximately constant and equal to p_1 , and the gradient of p_1 prevents motion parallel to the generators in a fluid of density ρ_1 (equation (8)); but the density varies rapidly across a section of the boundary layer, and inside the layer the pressure gradient does not balance the weight of the fluid. In addition to the transverse drift w_r , unless $l = 0$ the force of buoyancy sets up a convection current that contributes to the boundary-layer flow parallel to the main stream; when $l = 0$ the generators of the cylinder are vertical.

2.3. Primary and secondary flows

If gravity is neglected altogether, equation (1) and the modified equations (7) to (13) describe a flow in the boundary layer induced solely by, and nearly parallel to, the given main stream. This flow will be called the primary flow. The difference, due to gravity, between the actual flow and the primary flow will be termed secondary flow. The equations being non-linear, the secondary flow is not obtained correctly by neglecting the main stream and taking account of the gravity terms only.

The primary flow (gravity neglected) is independent of z as required, and satisfies the equations

$$\frac{\partial}{\partial x} (\rho_r u_r) + \frac{\partial}{\partial y} (\rho_r v_r) = 0, \quad (14)$$

$$\rho_r \left(u_r \frac{\partial u_r}{\partial x} + v_r \frac{\partial u_r}{\partial y} \right) - \rho_1 u_1 \frac{d u_1}{d x} = \frac{\partial}{\partial y} \left(\mu_r \frac{\partial u_r}{\partial y} \right), \quad (15)$$

$$\rho_r \left(u_r \frac{\partial i_r}{\partial x} + v_r \frac{\partial i_r}{\partial y} \right) - u_r \rho_1 \frac{d i_1}{d x} = \frac{1}{\sigma} \frac{\partial}{\partial y} \left(\mu_r \frac{\partial i_r}{\partial y} \right) + \mu_r \left(\frac{\partial u_r}{\partial y} \right)^2, \quad (16)$$

with
$$\frac{di_1}{dx} + u_1 \frac{du_1}{dx} = 0. \quad (17)$$

It is necessary now to discriminate between flows to which the primary flow alone provides a good approximation, and flows in which the secondary current must be taken into account. For simplicity the only cases considered in detail will be those in which the cylinder is either horizontal or vertical; an inclined cylinder will exhibit a combination of the features of these two extreme cases.

2.31. Horizontal cylinders

When the cylinder is horizontal ($n = 0$), the quantities p_1, ρ_1 , etc., are independent of z , and $w_r = 0$, even when gravity is taken into account. The exact boundary-layer equations for the total flow are (14) and (16) together with

$$\rho_r \left(u_r \frac{\partial u_r}{\partial x} + v_r \frac{\partial u_r}{\partial y} \right) - \rho_1 u_1 \frac{du_1}{dx} = g(\rho_1 - \rho_r)l + \frac{\partial}{\partial y} \left(\mu_r \frac{\partial u_r}{\partial y} \right), \quad (18)$$

and
$$\frac{di_1}{dx} + gl + u_1 \frac{du_1}{dx} = 0, \quad (19)$$

in place of (15) and (17). In this case the neglect of gravity amounts simply to omitting the gravity terms in (18) and (19), which is permissible if

$$u_1^2 \gg g \left| \int_0^x l dx \right|. \quad (20)$$

When (20) holds, the secondary flow is very much smaller than the primary flow, to which it is parallel. The condition (20) will hold when x is not too large, except perhaps near the foremost generator of a blunt-nosed cylinder, where u_1 is zero (see below).

If we consider a flat plate in a stream in its own plane, the above notation and equations apply when the z -axis is along the leading edge; and the plate is to be considered as a horizontal cylinder when the z -axis is horizontal, since the 'generators' are to be taken parallel to the leading edge. When the whole plane of the plate is horizontal, $l = n = 0$, and there is no secondary flow at all; the equations (14) to (17), with $u_1 = \text{const.}$, exactly describe the flow. If the plate is inclined at an angle α above the horizontal ($-\frac{1}{2}\pi \leq \alpha \leq \frac{1}{2}\pi$), then the condition (20) becomes

$$u_1^2 \gg gx \sin |\alpha|,$$

which is satisfied for a distance of, say, $u_1^2/(10g \sin |\alpha|)$ from the leading edge. A similar remark applies to any sharp-nosed cylinder, for which $u_1(0) \neq 0$; the secondary flow will be negligible for at least some distance along the boundary.

Consider next a horizontal blunt-nosed cylinder in a stream inclined at an angle α above the horizontal. Then $u_1(0) = 0$, and the application of the condition (20) near the foremost generator must be examined more closely. If the distance x is measured from the foremost generator along the upper side of the cylinder, then for small values of x ,

$$l \doteq \cos(\alpha - \kappa x),$$

where κ is the curvature of a normal cross-section at the front stagnation point, and with $u_1 \doteq \beta_1 x$ the condition (20) becomes

$$\beta_1^2 x^2 \gg \frac{g}{\kappa} |\sin \alpha - \sin(\alpha - \kappa x)|,$$

i.e.

$$\beta_1^2 x \gg g |\cos \alpha + \frac{1}{2} \kappa x \sin \alpha|.$$

If $\alpha = \pm \frac{1}{2}\pi$ this requires

$$\beta_1^2 x \gg g \cos \alpha,$$

which cannot be satisfied when x is small. If, however, $\alpha = \pm \frac{1}{2}\pi$, so that the cylinder is in a vertical stream, the condition (20) is fulfilled if

$$\beta_1^2 \gg \frac{1}{2} g \kappa. \quad (21)$$

Near the front stagnation point of a section of a horizontal blunt-nosed cylinder, therefore, the primary-flow equations are invalid except when the stream is vertical and the condition (21) is satisfied. Except in this one case the exact equations (14), (16), (18) and (19) are required. A simple solution of the exact equations has been found when $u_1 \propto x$, $l \propto x$ with dissipation in the energy equation neglected (§ 7), but since $l \propto x$, this helps only by providing the missing solutions for cylinders in a vertical stream for which the condition (21) is not satisfied. No solutions have been obtained for blunt horizontal cylinders in a non-vertical stream.

For a smooth cylinder β_1 will depend on κ , and the condition (21) needs further consideration. Evidence is provided by Ferri's (1942) experiments on a circular cylinder of diameter d at Mach numbers $M_1 \doteq 2$, for which $\beta_1/a_1 \doteq 3/d$, where a_1 is the velocity of sound in the incident undisturbed stream. If we accept that a similar order of magnitude is to be expected for any blunt-nosed cylinder at comparable Mach numbers, with d replaced by twice the radius of curvature at the nose, then $\beta_1 \doteq 1.5a_1\kappa$, and the condition (21) becomes

$$4.5\kappa \gg g/a_1^2.$$

This condition is satisfied unless the radius of curvature at the nose is very large.

2.32. Vertical cylinders

For a vertical cylinder, $l = 0$. We must now begin by neglecting the variations with z of the variables of state, p_1 , ρ_1 and i_1 , in the main stream, which will give a satisfactory approximation if conditions are being studied on a local rather than on a meteorological scale, as they usually are in boundary-layer theory. Then none of the dependent variables will depend on z . The equation of continuity reduces to (14), and, since $l = 0$, the general equation (11) reduces to (15), and the general equation (7) to (17). There will be a cross-drift w_r , but in the general equation (13), assuming that w_r is small, the terms $w_r(\partial p_1/\partial z)$ and $\mu_r(\partial w_r/\partial y)^2$ will both be negligibly small. It will be shown in § 6 that their omission is permissible not only for a flat plate and for a sharp-nosed cylinder but also for a blunt-nosed cylinder. The result is that (13) also reduces to (16), so the flow in the x -direction is very nearly the primary flow obtained from equations (14) to (17). Equation (12) now becomes

$$\rho_r \left(u_r \frac{\partial w_r}{\partial x} + v_r \frac{\partial w_r}{\partial y} \right) = g(\rho_1 - \rho_r) + \frac{\partial}{\partial y} \left(\mu_r \frac{\partial w_r}{\partial y} \right); \quad (22)$$

it is independent of the other equations and may be solved for w_r (see § 6).

2.33. The transformed equations of primary flow

The position is that the primary flow is a good approximation to the total flow not too far from the leading edge of a flat plate or a sharp-nosed cylinder whether the leading edge is horizontal or vertical, and therefore probably also for any inclination of the leading edge. For a blunt-nosed cylinder the same result holds when the cylinder is vertical in a horizontal stream, and when the cylinder, provided it is not too blunt, is horizontal (and probably also whatever its position) in a vertical stream. The equations (14) to (17) of primary flow are discussed in the following three sections, for a flat plate in § 3, near the front stagnation point of a section of a blunt-nosed cylinder in § 4, and for any cylinder—i.e. for any given distribution of main-stream velocity—in § 5.

Before the equations are discussed, they are transformed by Mises's transformation. Equation (14) is replaced by

$$\rho_r u_r = \partial\psi/\partial y, \quad \rho_r v_r = -\partial\psi/\partial x, \quad (23)$$

where $\psi(x, y)$ is the stream function. The independent variables are then changed from x, y to x, ψ , whereby $\partial/\partial x$ is replaced by $\partial/\partial x - \rho_r v_r \partial/\partial\psi$ and $\partial/\partial y$ by $\rho_r u_r \partial/\partial\psi$. Non-dimensional functions u, ρ , etc., defined by $u_r = u_1 u$, $\rho_r = \rho_1 \rho$, etc., are also introduced, together with $\phi_1 = \mu_1 \rho_1 u_1$ and $\phi = \mu \rho u$, and in place of x the co-ordinate $s = \int_0^x \phi_1(\tau) d\tau$ is used. Then the equations (15) and (16) become

$$\left(\frac{d}{ds} \ln u_1\right)(u^2 - T) + u \frac{\partial u}{\partial s} = u \frac{\partial}{\partial\psi} \left(\phi \frac{\partial u}{\partial\psi}\right), \quad (24)$$

$$\frac{di_1}{ds}(i - T) + i_1 \frac{\partial i}{\partial s} = \frac{i_1}{\sigma} \frac{\partial}{\partial\psi} \left(\phi \frac{\partial i}{\partial\psi}\right) + u_1^2 \phi \left(\frac{\partial u}{\partial\psi}\right)^2. \quad (25)$$

Since $p_r = p_1$, in the new notation $\rho T = 1$, from the equation of state.

3. THE BOUNDARY LAYER OF A FLAT PLATE ALONG A STREAM

The main stream past a plate occupying the half-plane $y = 0, x \geq 0$ is taken as uniform. For a horizontal plate there is (accurately) no secondary flow, and the equations for the boundary-layer flow are

$$\frac{\partial u}{\partial s} = \frac{\partial}{\partial\psi} \left(\phi \frac{\partial u}{\partial\psi}\right), \quad (26)$$

$$i_1 \frac{\partial i}{\partial s} = \frac{i_1}{\sigma} \frac{\partial}{\partial\psi} \left(\phi \frac{\partial i}{\partial\psi}\right) + u_1^2 \phi \left(\frac{\partial u}{\partial\psi}\right)^2, \quad (27)$$

with the following boundary conditions:

$$u = 0 \quad \text{when} \quad \psi = 0, s > 0; \quad u \rightarrow 1, i \rightarrow 1 \quad \text{as} \quad \psi \rightarrow \infty$$

for all values of s ; $u = 1, i = 1$ when $s = 0$ for all values of ψ except 0; and a condition on i when $\psi = 0, s > 0$ that specifies the thermal condition of the plate (see below).

For a plate in any position the secondary flow may be neglected for a certain distance along the plate from the leading edge, and the primary flow is described by equations (26) and (27).

These equations and the boundary conditions can all be satisfied with u , i and ϕ functions of the single variable $\eta = \psi/(2s)^{\frac{1}{2}}$ only; with η as independent variable (26) and (27) reduce to the ordinary differential equations

$$\frac{d}{d\eta} \left(\phi \frac{du}{d\eta} \right) + \eta \frac{du}{d\eta} = 0, \quad (28)$$

$$\frac{d}{d\eta} \left(\phi \frac{di}{d\eta} \right) + \sigma \eta \frac{di}{d\eta} + \sigma b \phi \left(\frac{du}{d\eta} \right)^2 = 0, \quad (29)$$

where $b = u_1^2/i_1 = (\gamma_1 - 1) M_1^2$, M_1 being the Mach number of the main stream.

The independent variable is now changed from η to ζ , where†

$$\zeta = \int_0^\eta \frac{d\tau}{u(\tau)},$$

so that $u = d\eta/d\zeta$, and (28) and (29) become

$$\frac{d}{d\zeta} (\theta \eta'') + \eta \eta'' = 0, \quad (30)$$

$$\frac{d}{d\zeta} (\theta i') + \sigma \eta i' + \sigma b \theta \eta'^2 = 0, \quad (31)$$

where $\theta = \mu\rho$ (so that $\phi = \theta u$) and dashes denote differentiation with respect to ζ . The boundary conditions become $\eta(0) = \eta'(\infty) = 0$; $\eta'(\infty) = i(\infty) = 1$, together with a condition on i at $\zeta = 0$ (see below).

In these equations θ is a function of i only. For μ_r is a function of T_r only, so $\mu\rho = \mu/T$ is a function of T only. Also since i_r is an increasing function of T_r , T is a known function of i ; indeed, when the specific heats may be regarded as constant, $T = i$. A discussion of the function is given by Crocco (1946). He takes Sutherland's formula

$$\mu_r = \mu_r(T_r) = \mu_r(273) \left(\frac{T_r}{273} \right)^{\frac{1}{2}} \frac{1 + T_c/273}{1 + T_c/T_r},$$

where T_c is a characteristic temperature depending on the gas, as the most accurate representation of the dependence of μ_r on T_r . This formula gives

$$\theta = \sqrt{T} \frac{1 + T^*}{T + T^*}, \quad (32)$$

where $T^* = T_c/T_1$. A formula that is sometimes more convenient for computation is

$$\mu_r(T_r) = \mu_r(273) \left(\frac{T_r}{273} \right)^{\epsilon+1},$$

which implies that

$$\theta = T^\epsilon. \quad (33)$$

This formula, with $\epsilon = -0.24$ for air, has been used by previous authors (Kármán & Tsien 1938). Crocco examines to what extent (33) can be considered as a good approximation to (32). He shows that for a boundary layer in which the temperature

† $u = O(\eta^{\frac{1}{2}})$ as $\eta \rightarrow 0$, and the integral converges.

does not vary much, so that $T \doteq 1$, the best agreement between the two formulae is obtained by taking $\epsilon = \frac{1}{2} \frac{T^* - 1}{T^* + 1}$ in (33). Corresponding numerical values of ϵ are given in the following table (extracted from Crocco's paper):

	σ	T_1 ($^{\circ}$ K)	T^*	ϵ
air	0.725	116	1	0
		300	0.387	-0.221
		3000	0.039	-0.463
		300	2.24	0.192
water vapour	1.08	673	1	0
		3000	0.224	-0.317

The case $\epsilon = 0$ has special importance, since then $\mu_r \propto T_r$ and $\theta = 1$, which leads to considerable simplification of the equations.

After this digression on the variation of θ with i , we return to the integration of (30) and (31) for *any* known or assumed law of variation. In the first place, for $\sigma = 1$ the solution of (30) and (31) has been given by Kármán & Tsien (1938). When $\sigma = 1$ the energy equation (31) is equivalent to the momentum equation (30) provided that

$$i = i(0) + [1 - i(0) + \frac{1}{2}b]\eta' - \frac{1}{2}b\eta'^2, \quad (34)$$

where $i(0) = i_0/i_1$, i_0 being the enthalpy at the plate. Since θ is a function of i only, (34) can be used when $\sigma = 1$ to express θ in terms of η , and then (30) can be solved for the one dependent variable η . If, however, $\sigma \neq 1$, (30) and (31) must be solved simultaneously for η and i . A solution, based on different equations, has been given by Crocco (1946), so it will suffice to indicate how Kármán & Tsien's method of integration may be adapted in this more general case.

Equations (30) and (31), with the relevant boundary conditions, may be transformed into the integral equations

$$u(\zeta) = \eta' = F(\zeta)/F(\infty), \quad (35)$$

$$i(\zeta) = i(0) + \theta(0) i'(0) G(\zeta) - \sigma b H(\zeta), \quad (36)$$

where the new functions are defined by

$$\left. \begin{aligned} E(\zeta) &= \exp\left(-\int_0^\zeta \frac{\eta}{\theta} d\zeta\right), & F(\zeta) &= \int_0^\zeta \frac{E}{\theta} d\zeta, \\ G(\zeta) &= \int_0^\zeta \frac{E^\sigma}{\theta} d\zeta, & H(\zeta) &= \frac{1}{F^2(\infty)} \int_0^\zeta \frac{E^\sigma}{\theta} \left(\int_0^\zeta \frac{E^{2-\sigma}}{\theta} d\zeta\right) d\zeta. \end{aligned} \right\} \quad (37)$$

The constant $\theta(0) i'(0)$ in (36) can be expressed in terms of the rate of transfer of heat from the plate to the gas. For the rate of transfer of heat across the plane $y = \text{const.}$ is

$$q_r = -\frac{\mu_r}{\sigma} \frac{\partial i_r}{\partial y} = -\frac{i_1}{\sigma} \left(\frac{\phi_1}{2x}\right)^{\frac{1}{2}} \theta i'$$

per unit area, so the rate of transfer from a strip of the plate of length L and unit breadth is

$$Q_0 = \int_0^L q_0 dx = -\frac{i_1}{\sigma} (2\phi_1 L)^{\frac{1}{2}} \theta(0) i'(0),$$

where the suffix 0 refers, as usual, to conditions at the boundary. To specify the thermal condition of the plate either Q_0 or $i(0)$ may be given, and the other is then found from the equation

$$1 = i(0) - \frac{\sigma Q_0}{i_1(2\phi_1 L)^{\frac{1}{2}}} G(\infty) - \sigma b H(\infty). \quad (38)$$

The solution is found by successive approximation, starting from the solution when $\theta = 1$. When $\theta = 1$, (30) and (31) are independent of each other, and both equations occur in the theory of the flow of an incompressible fluid. In fact, (30) is then Blasius's equation

$$\eta''' + \eta\eta'' = 0, \quad (39)$$

and the transformation $i(\zeta) = 1 + \frac{1}{2}b\vartheta(\zeta)$ changes (31) into

$$\vartheta'' + \sigma\eta\vartheta' + 2\sigma\eta'' = 0, \quad (40)$$

which has been discussed by Pohlhausen (1921).

When a more accurate expression for θ in terms of i is used, the values of i from the known solution of (40) lead to a first approximation to θ , which, in conjunction with the known solution of (39) for η , can be used to provide first approximations to E , F , G and H . Formulae (35), (36) and (38) then give improved values for η , u and i , and the process is repeated.

4. THE PRIMARY FLOW IN THE BOUNDARY LAYER NEAR THE FOREMOST GENERATOR OF A BLUNT-NOSED CYLINDER

The solution given in this section is a good approximation for a vertical cylinder in a horizontal stream (as will be shown in § 6), and for a horizontal cylinder in a vertical stream for which the condition (21) is satisfied. The approximate considerations of §§ 2 and 6 suggest that for a sloping cylinder the primary-flow solution is a good approximation under the same conditions as for a horizontal cylinder.

In general, with u_1 and i_1 not uniform, the equations (24) and (25) are not easily solved. If, however, (i) c_p is constant, so that $i = T$, (ii) the dissipation is negligible in (25), and (iii) $u_1 = \beta s^m$, where β is a constant, then the equations are reducible to the ordinary differential equations

$$\frac{d}{d\zeta}(\theta\eta'') + \eta\eta'' - 2m(\eta'^2 - i) = 0, \quad (41)$$

$$\frac{d}{d\zeta}(\theta i') + \sigma\eta i' = 0, \quad (42)$$

where $\eta = \frac{\psi}{(2s)^{\frac{1}{2}}}$, $\zeta = \int_0^\eta \frac{d\tau}{u(\tau)}$ as before, and η is regarded as a function of ζ with $u = d\eta/d\zeta$. The boundary conditions are the same as for equations (26) and (27).

At the front of a blunt-nosed cylinder the main stream is brought to rest, and its velocity then increases proportionally to x , so the equations are appropriate near the front of a blunt-nosed cylinder only if $u_1 \propto x$. If the suffix 10 is used to refer to the foremost generator of the cylinder, then from Bernoulli's equation

$$\frac{T_1}{T_{10}} = 1 - \frac{u_1^2}{2c_p T_{10}} = 1 - \frac{\gamma - 1}{2} \frac{u_1^2}{a_{10}^2},$$

when c_p is constant. Near the front of the cylinder

$$\phi_1 = \theta_{10} u_1 + O(u_1^3),$$

so

$$x = \frac{s^{1-m}}{\theta_{10} \beta (1-m)}$$

and

$$u_1 = \beta_1 x^{m/(1-m)} \quad \text{approximately,}$$

where $\beta_1 = \beta[\theta_{10}\beta(1-m)]^{m/(1-m)}$. It follows that the solution of (41) and (42), with $m = \frac{1}{2}$, will describe the primary flow near the foremost generator of a blunt-nosed cylinder provided that the heat of dissipation is negligible. As for incompressible fluids (see Goldstein 1938, pp. 608, 609) the dissipation is negligible if $u_1^2/|i_1 - i_0|$ is small; if $u_1^2/|i_1 - i_0|$ is not to exceed 1/10, for example, then in air u_1 must not exceed $33\Delta T_r$ ft./sec., ΔT_r °K being the difference between the temperature of the cylinder and the stagnation temperature of the main stream.

$$\text{With } m = \frac{1}{2}, \quad u_1 = \beta_1 x, \quad s = \frac{1}{2}\theta_{10}\beta_1 x^2, \quad \text{and} \quad \eta = \frac{\psi}{x(\theta_{10}\beta_1)^{\frac{1}{2}}},$$

equations (41) and (42), with the appropriate boundary conditions, are transformed into the integral equations

$$u(\zeta) = \frac{1 - I(\infty)}{F(\infty)} F(\zeta) + I(\zeta), \quad (43)$$

and

$$i(\zeta) = i(0) + \theta(0) i'(0) G(\zeta), \quad (44)$$

with the condition

$$1 = i(0) + \theta(0) i'(0) G(\infty),$$

where

$$I(\zeta) = \int_0^\zeta \frac{E}{\theta} \left(\int_0^\zeta \frac{\eta'^2 - i}{E} d\zeta \right) d\zeta,$$

and E , F and G were defined in equation (37).

There is an obstacle to the computation of I because $\eta'^2 - i$ and E both tend to zero as $\zeta \rightarrow \infty$, and in the one calculation that was made (43) was used only to provide an estimate of the values of u . As a first approximation, it is more convenient to start with the known solution of (41) and (42) for an incompressible fluid than with the solution for $\theta = 1$. The equations for an incompressible fluid are

$$\eta''' + \eta\eta'' - \eta'^2 + 1 = 0, \quad \vartheta'' + \sigma\eta\vartheta' = 0, \quad (45)$$

where

$$i(\zeta) = i(0) + [1 - i(0)] \vartheta(\zeta),$$

and the boundary conditions are

$$\eta(0) = \eta'(0) = \vartheta(0) = 0, \quad \eta'(\infty) = \vartheta(\infty) = 1.$$

[The equations for a compressible fluid with $\theta = 1$ differ from (45) only in having i instead of 1 in the first equation.] The values of $u(\zeta)$ and $i(\zeta)$ for an incompressible fluid are easily found for any given $i(0)$.

Equations (41) (with $m = \frac{1}{2}$) and (42) were solved with $\sigma = 0.725$, and $i(0) = 0.8$, which corresponds to the flow of air at the speed of sound against a circular cylinder that is maintained at the temperature of the undisturbed part of the incident stream. Sutherland's law was used for the variation of viscosity with enthalpy. Starting

with the solution of (45), approximate values of u were obtained from (43) and, at the same time, values of i which proved to be correct to three decimal places, in spite of the inaccuracies in u , were obtained from (44). Final values of u were obtained by integrating (41) directly. With i , and therefore θ , determined, equation (41) involves only the single unknown function η , and was integrated, in the same way as the first of equations (45) (see Howarth 1934), by straddling the true solution (for which $\eta'(\infty) = 1$) by solutions with $\eta(0) = \eta'(0) = 0$ and with estimated high and low values of $\eta''(0)$, and finally interpolating. It was, of course, necessary finally to use (44) again, to verify that i had been determined sufficiently accurately in the first stage.

Values comparing the solutions for an incompressible fluid and for air are given in table 1. The Nusselt number is

$$Nu = \left(\frac{\beta_1 d^2}{\nu_{10}} \right)^{\frac{1}{2}} / G(\infty),$$

where d is a representative length. To the present accuracy it is the same as for an incompressible fluid with $\sigma = 0.725$, viz.

$$Nu = 0.502 \left(\frac{\beta_1 d^2}{\nu_{10}} \right)^{\frac{1}{2}}.$$

TABLE 1

ζ	incompressible fluid				air				
	η	$\eta' = u$	η''	i	η	$\eta' = u$	η''	i	w (see § 6)
0	0	0	1.2326	0.800	0	0	1.0838	0.800	0
0.2	0.023	0.223	1.034	0.820	0.021	0.202	0.936	0.819	0.06
0.4	0.088	0.414	0.846	0.840	0.079	0.374	0.790	0.838	0.11
0.6	0.187	0.566	0.675	0.860	0.168	0.519	0.653	0.858	0.14
0.8	0.312	0.686	0.525	0.879	0.284	0.636	0.528	0.876	0.16
1.0	0.459	0.778	0.398	0.898	0.422	0.731	0.418	0.894	0.16
1.2	0.622	0.847	0.294	0.915	0.575	0.805	0.323	0.912	0.15
1.4	0.797	0.897	0.211	0.930	0.742	0.861	0.244	0.928	0.14
1.6	0.980	0.932	0.147	0.944	0.919	0.903	0.180	0.942	0.12
1.8	1.169	0.957	0.100	0.957	1.103	0.934	0.130	0.954	0.10
2.0	1.362	0.973	0.066	0.967	1.292	0.956	0.092	0.965	0.08
2.5	1.854	0.993	0.020	0.985	1.779	0.986	0.034	0.985	0.04
3.0	2.352	0.998	0.005	0.994	2.274	0.996	0.011	0.995	0.01
3.5	2.852	1.000	0.001	0.998	2.774	1.000	0.003	0.999	0.00
4.0	3.352	1.000	0.000	0.999	3.273	1.000	0.000	1.000	0

5. THE PRIMARY FLOW IN THE BOUNDARY LAYER OF A CYLINDER

The solution discussed in § 4 is valid only in the vicinity of the foremost generator of the cylinder, because it involves the assumption that u_1 is small enough for dissipation to be negligible and is expressible in the form $u_1 = \beta s^{\frac{1}{2}} \doteq \beta_1 x^{\frac{1}{2}}$. The question arises whether the primary flow in the whole laminar boundary layer of a cylinder (and therefore, when the secondary flow is negligible, a good approximation to the total flow) can be calculated from ordinary differential equations. For this purpose a single term $\beta s^{\frac{1}{2}}$ will be inadequate to represent u_1 everywhere,

though it may be represented by a series of powers of s (or x). There is a corresponding situation in the case of an incompressible fluid; when the main stream is proportional to x the solution involves the integration of one ordinary differential equation (Howarth 1934), but when the main stream is represented by a series of powers of x the solution involves the successive integration of a number of ordinary differential equations, only the first being non-linear (see Howarth 1934, 1938). Similarly, for a compressible fluid, when $u_1 = \sum_{n=0} b_n x^n$ the solution involves a number of ordinary differential equations. Such a solution for an accelerated main stream has been given by Howarth (1948), and for a retarded main stream by Cope & Hartree (1948) and by Howarth (1948). The object of mentioning these problems again is to show that under the restrictions imposed by Howarth ($\sigma = 1$, $\theta = 1$, heat-insulated boundary) solutions may be obtained from solutions for an incompressible fluid without introducing any new differential equations.

If c_p is constant so that $i = T$, (24) and (25) may be written as

$$\left(\frac{d}{ds} \ln u_1\right)(u^2 - i) + u \frac{\partial u}{\partial s} = u \frac{\partial}{\partial \psi} \left(\phi \frac{\partial u}{\partial \psi}\right), \quad (46)$$

$$\frac{\partial i}{\partial s} = \frac{1}{\sigma} \frac{\partial}{\partial \psi} \left(\phi \frac{\partial i}{\partial \psi}\right) + b \phi \left(\frac{\partial u}{\partial \psi}\right)^2, \quad (47)$$

where $s = \int_0^x \phi_1 dx$ and $b = u_1^2/i_1 = (\gamma - 1) M_1^2$, as before. For an incompressible fluid of uniform density ρ_{i1} the momentum equation corresponding to (46) is

$$\left(\frac{d}{ds} \ln u_{i1}\right)(u_i^2 - 1) + u_i \frac{\partial u_i}{\partial s} = u_i \frac{\partial}{\partial \psi} \left(u_i \frac{\partial u_i}{\partial \psi}\right), \quad (48)$$

where the suffix i is used to characterize the incompressible fluid, and

$$s = \theta_{i1} \int_0^{x_i} u_{i1} dx_i, \quad \rho_{i1} u_{ir} = \frac{\partial \psi}{\partial y_i}, \quad \rho_{i1} v_{ir} = -\frac{\partial \psi}{\partial x_i}.$$

If $\mu_r \propto T_r$, then $\theta = 1$ and $\phi = u$. If also $\sigma = 1$ and the boundary is heat-insulated, then, use being made of the equation $i_1 + \frac{1}{2}u_1^2 = \text{const.}$, it may be proved that the relevant solution of (46) and (47) is provided by

$$i + \frac{1}{2}bu^2 = 1 + \frac{1}{2}b,$$

together with the relevant solution of

$$\left[(1 + \frac{1}{2}b) \frac{d}{ds} \ln u_1\right][u^2 - 1] + u \frac{\partial u}{\partial s} = u \frac{\partial}{\partial \psi} \left(u \frac{\partial u}{\partial \psi}\right). \quad (49)$$

In terms of s and ψ , the equation (48) satisfied by u_i and the equation (49) satisfied by u are then the same if

$$(1 + \frac{1}{2}b) \frac{d}{ds} \ln u_1 = \frac{d}{ds} \ln u_{i1}, \quad (50)$$

and the boundary conditions are the same for both equations. Hence, corresponding points (x, y) and (x_i, y_i) in the compressible and incompressible fluids being defined as points for which s and ψ have the same values in the two cases, it follows that

the values of $u(x, y)$ and $u_i(x_i, y_i)$ at corresponding points are the same, if (50) is satisfied. The integral of (50) is

$$M_1 = \frac{u_{i1}}{A}, \quad (51)$$

where A is an arbitrary constant. Equation (51) must hold at points corresponding to the same value of s in the two fluids, i.e. if x and x_i correspond, in the compressible and incompressible fluids, respectively, to the same value of s , $u_{i1}(x_i) = AM_1(x)$. Now, for a given problem in a compressible fluid, $M_1(x)$ will be known, and therefore (51) fixes the incompressible flow problem that should be solved, once A has been chosen and the relation between x_i and x are worked out. The relation between x_i and x is found from

$$\frac{dx_i}{dx} = \left(\frac{ds}{dx} \right) / \left(\frac{ds}{dx_i} \right) = \frac{\phi_1}{\phi_{i1}}.$$

Since $\mu_r \propto T_r$, $\theta_1 = \theta_{10} p_1 / p_{10} = \theta_{10} (i_1 / i_{10})^{\gamma/(\gamma-1)}$, where the suffix 10 again refers to the stagnation point. It follows that

$$\frac{dx_i}{dx} = \frac{\theta_{10} a_{10}}{\theta_{i1} A} \left(\frac{i_1}{i_{10}} \right)^{\gamma/(\gamma-1)} \frac{a_1}{a_{10}} = \frac{\theta_{10} a_{10}}{\theta_{i1} A} \left(\frac{T_1}{T_{10}} \right)^{(3\gamma-1)/(2(\gamma-1))} = \frac{\theta_{10} a_{10}}{\theta_{i1} A} (1 + \frac{1}{2}b)^{-(3\gamma-1)/(2(\gamma-1))}.$$

It will be assumed, without loss of generality, that $A = a_{10}$ and that the density and viscosity of the incompressible fluid are the same as those of the compressible fluid at the stagnation point, so that $\theta_{i1} = \theta_{10}$. Then

$$x_i = \int_0^x [1 + \frac{1}{2}b(x)]^{-(3\gamma-1)/(2(\gamma-1))} dx. \quad (52)$$

With this relation between x_i and x , the flow of the incompressible fluid in a boundary layer is then determined for the main-stream velocity distribution

$$u_{i1}(x_i) = a_{10} M_1(x), \quad (53)$$

and the solution of the compressible flow problem is provided by $u(x, y) = u_i(x_i, y_i)$. The relation between y and y_i is

$$\begin{aligned} y &= \frac{1}{\rho_1 u_1} \int_0^y \frac{d\psi}{\rho u} = \frac{1}{\rho_1 u_1} \int_0^y \frac{i}{u_i} d\psi = \frac{\rho_{i1} u_{i1}(x_i)}{\rho_1(x) u_1(x)} \int_0^{y_i} i dy_i \\ &= [1 + \frac{1}{2}b(x)]^{(\gamma+1)/(2(\gamma-1))} \int_0^{y_i} \{1 + \frac{1}{2}b(x) [1 - u_i^2(x_i, y_i)]\} dy_i. \end{aligned} \quad (54)$$

For example, when the main stream is the accelerated flow away from the foremost generator of a blunt-nosed cylinder and the main-stream Mach number is given by

$$M_1(x) = \sum_{n=1} b_n x^n, \quad (55)$$

where the b_n 's are given constants, then from (52), (53) and (54) the flow in the boundary layer may be obtained from the flow of an incompressible fluid (see Howarth 1934) in the boundary layer near the front of a cylinder for which

$$u_{i1}(x_i) = \sum_{n=1} b_{in} x_i^n, \quad (56)$$

where the b_{in} 's are known constants.

In § 4 an approximate solution was found for $u_1 = \beta s^m$ for any value of σ but with dissipation neglected, and the neglect of dissipation is not permissible at a heat-insulated wall. By the present method a solution can be found for $u_1 = \beta s^m$ when $\sigma = 1$ and the wall is heat-insulated.

$$\text{From (51)} \quad u_{i1} = \frac{u_1}{(1 - \frac{1}{2}u_1^2/i_{10})^{\frac{1}{2}}},$$

$$\begin{aligned} \text{and} \quad x_i &= \frac{1}{\theta_{10}} \int_0^s \frac{ds}{u_{i1}} = \frac{1}{m\theta_{10}\beta^{1/m}} \int_0^{u_1} u_1^{1/m-2} \left(1 - \frac{u_1^2}{2i_{10}}\right)^{\frac{1}{2}} du_1 \\ &= \frac{1}{m\theta_{10}\beta^{1/m}} \int_0^{u_{i1}} \frac{u_{i1}^{1/m-2}}{\left(1 + \frac{u_{i1}^2}{2i_{10}}\right)^{1/(2m)+1}} du_{i1}, \end{aligned}$$

for $m < 1$, whence u_{i1} is determined in terms of x_i . In particular, for $m = \frac{1}{2}$,

$$x_i = \frac{(2i_{10})^{\frac{1}{2}}}{\theta_{10}\beta^2} \left(\tan^{-1} t + \frac{t}{1+t^2} \right),$$

$$\text{where} \quad t^2 = \frac{u_{i1}^2}{2i_{10}},$$

$$\text{i.e.} \quad x_i = 2 \frac{(2i_{10})^{\frac{1}{2}}}{\theta_{10}\beta^2} \left(t - \frac{2}{3}t^3 + \frac{2}{5}t^5 - \dots \right),$$

and, by inverting the series,

$$u_{i1} = (2i_{10})^{\frac{1}{2}} \left[\alpha x_i + \frac{2}{3}(\alpha x_i)^3 + \frac{1}{15}(\alpha x_i)^5 + \dots \right],$$

where $\alpha = \frac{\theta_{10}\beta^2}{2(2i_{10})^{\frac{1}{2}}}$. This may now be considered a particular case of the preceding example (see (56)).

For the first of the two examples mentioned above, and for the second with $m = \frac{1}{2}$, the primary flow is a good approximation to the total flow under the conditions set out at the beginning of § 4.

As an example of retarded flow, we may employ the reverse procedure and start from the main stream of an incompressible fluid

$$u_{i1}(x_i) = b_{i0} - b_{i1}x_i, \quad (57)$$

where b_{i0} , b_{i1} are positive constants. The velocity distribution function $u_i(x_i, y_i)$ has been calculated by Howarth (1938) and Hartree (1949).

$$\text{The relation} \quad x = \int_0^{x_i} \left[1 + \frac{\gamma-1}{2} \left(\frac{b_{i0} - b_{i1}x_i}{a_{10}} \right)^2 \right]^{(3\gamma-1)/(2(\gamma-1))} dx_i,$$

together with (54) and (57), enables M_1 to be expressed as a function of x . The distribution of main-stream velocity in the compressible fluid is complicated, but for small enough values of x

$$x \doteq Bx_i, \quad \text{where} \quad B = \left[1 + \frac{\gamma-1}{2} \left(\frac{b_{i0}}{a_{10}} \right)^2 \right]^{(3\gamma-1)/(2(\gamma-1))},$$

$$\text{so} \quad M_1(x) = \frac{1}{a_{10}} \left(b_{i0} - \frac{b_{i1}}{B} x \right) + O(x^2).$$

The solution therefore applies to a non-uniform flow past a sharp-nosed cylinder with retardation in the vicinity of the foremost generator. The primary flow will be a good approximation to the total flow whatever the orientation of the cylinder and whatever the direction of the stream (see § 2).

If $\sigma \neq 1$ (and even with $\sigma = 1$, if $\theta \neq 1$ or if there is heat transfer at the boundary) solutions of (46) and (47) cannot be readily found from solutions for incompressible flow, and it is necessary to obtain new solutions. Consider the case $\theta = i^\circ$. The main stream is given, so $d(\ln u_1)/ds$ and b can be expressed as power series

$$\frac{d}{ds} \ln u_1 = \sum_{n=0} c_n s^n, \quad b = \sum_{n=0} d_n s^n,$$

the c_n 's and d_n 's being known constants. Write

$$u = \sum_{n=0} f_n(\eta) s^n, \quad i = \sum_{n=0} g_n(\eta) s^n,$$

where $\eta = \psi/(2s)^{\frac{1}{2}}$ as before. Then

$$\frac{d}{d\eta} \left(g_0^e f_0 \frac{df_0}{d\eta} \right) + \eta \frac{df_0}{d\eta} = 0,$$

$$\frac{1}{\sigma} \frac{d}{d\eta} \left(g_0^e f_0 \frac{dg_0}{d\eta} \right) + \eta \frac{dg_0}{d\eta} + d_0 g_0^e f_0 \left(\frac{df_0}{d\eta} \right)^2 = 0.$$

Now change the independent variable from η to $\zeta = \int_0^\eta \frac{d\eta}{f_0}$. These two equations become the same as (30) and (31) with g_0 in place of i and d_0 in place of (the constant) b in (31), and with $\theta = g_0^e$. The boundary conditions are the same, so η and g_0 are found as indicated in § 3. The equations for f_n , g_n ($n \geq 1$) are second-order, linear differential equations, with the boundary conditions

$$f_n(0) = g_n(0) = f_n(\infty) = g_n(\infty) = 0.$$

The first four are

$$\frac{d}{d\zeta} \left[g_0^e \left(f_1' + \frac{\eta''}{\eta'} f_1 \right) + \epsilon \eta'' g_0^{e-1} g_1 \right] + \eta f_1' - 2\eta' f_1 = 2c_0(\eta'^2 - g_0),$$

$$\frac{1}{\sigma} \frac{d}{d\zeta} \left(\frac{1}{\eta'} g_0^e g_0' f_1 + g_0^e g_1' + \epsilon g_0^{e-1} g_0' g_1 \right) + \eta g_1' - 2\eta' g_1 + d_0 \left(2\eta'' g_0^e f_1' + \frac{\eta''^2}{\eta'} g_0^e f_1 + \epsilon \eta''^2 g_0^{e-1} g_1 \right)$$

$$= -d_1 \eta''^2 g_0^e,$$

$$\frac{d}{d\zeta} \left[g_0^e \left(f_2' + \frac{\eta''}{\eta'} f_2 \right) + \epsilon \eta'' g_0^{e-1} g_2 \right] + \eta f_2' - 4\eta' f_2$$

$$= 2c_0(\eta' f_1 + \frac{g_0}{\eta'} f_1 - g_1) + 2c_1(\eta'^2 - g_0)$$

$$- \frac{d}{d\zeta} \left[\frac{1}{\eta'} g_0^e f_1 f_1' + \epsilon g_0^{e-1} \left(f_1' + \frac{\eta''}{\eta'} f_1 \right) g_1 + \frac{\epsilon(\epsilon-1)}{2} \eta'' g_0^{e-2} g_1^2 \right],$$

$$\begin{aligned}
& \frac{1}{\sigma} \frac{d}{d\zeta} \left(\frac{1}{\eta'} g_0^e g_0' f_2 + g_0^e g_2' + \epsilon g_0^{e-1} g_0' g_2 \right) + \eta g_2' - 4\eta' g_2 + d_0 \left(2\eta'' g_0^e f_2' + \frac{\eta''^2}{\eta'} g_0^e f_2 + \epsilon \eta''^2 g_0^{e-1} g_2 \right) \\
&= -d_0 \left[g_0^e f_1'^2 + \frac{2\eta''}{\eta'} g_0^e f_1 f_1' + 2\epsilon \eta'' g_0^{e-1} f_1' g_1 + \frac{\epsilon \eta''^2}{\eta'} g_0^{e-1} f_1 g_1 + \frac{\epsilon(\epsilon-1)}{2} \eta''^2 g_0^{e-2} g_1^2 \right] \\
&\quad - d_1 \left(2\eta'' g_0^e f_1' + \frac{\eta''^2}{\eta_1} g_0^e f_1 + \epsilon \eta''^2 g_0^{e-1} g_1 \right) - d_2 \eta''^2 g_0^e \\
&\quad - \frac{1}{\sigma} \frac{d}{d\zeta} \left[\frac{1}{\eta'} g_0^e f_1 g_1' + \frac{\epsilon}{\eta'} g_0^{e-1} g_0' f_1 g_1 + \epsilon g_0^{e-1} g_1 g_1' + \frac{\epsilon(\epsilon-1)}{2} g_0^{e-2} g_0' g_1^2 \right].
\end{aligned}$$

The first pair of these equations have to be solved simultaneously for f_1 , g_1 , and the second pair for f_2 , g_2 , and so on. A system of equations for this problem, but with quantities proportional to x and $\frac{1}{\sqrt{x}} \int_0^y \frac{dy}{\mu}$ as independent variables, is considered by Cope & Hartree (1948). Use of s and ζ has the advantage that only two simultaneous linear ordinary differential equations have to be considered at each stage, whereas three such equations are involved at each stage in Cope & Hartree's method. On the other hand, since η''/η' behaves like $1/\zeta$ near $\zeta = 0$ it is necessary to *start* the integration of any pair of the above equations at $\zeta = 0$ by expressing each of the dependent variables as a series of powers of ζ —a procedure that is not necessary in Cope & Hartree's method, where the origin is a regular point of the equations. In the particular case $\theta = 1$ ($\epsilon = 0$), the first of each pair of equations is independent of the second and the equations can be solved successively in the order given. There is a corresponding simplification when $\theta = 1$ in Cope & Hartree's equations; the three simultaneous equations that occur at each stage may be reduced to two, of which the first is independent of the second.

6. SECONDARY FLOW, ESPECIALLY FOR A VERTICAL CYLINDER

An approximate method of finding the drift w_r from equation (22) when the cylinder is vertical was explained in § 2. Equation (22) is transformed into

$$u_r \frac{\partial w_r}{\partial x} = g(T-1) + u_r \frac{\partial}{\partial \psi} \left(\phi_r \frac{\partial w_r}{\partial \psi} \right). \quad (58)$$

Solutions have been computed for a flat plate and for a blunt-nosed cylinder, for which simple solutions of the horizontal primary flow were given in §§ 3 and 4.

Consider first a flat plate, with its *leading edge vertical*, in a horizontal stream. In this case $u_1 = \text{const.}$, $\phi_1 = \text{const.}$, $u_r = u_1 \eta'(\zeta)$, $i_r = i_1 i(\zeta)$, where $\eta = \psi / (2\phi_1 x)^{\frac{1}{2}}$, and η and i are the solutions of (30) and (31). If c_p is constant, so that $i = T$, and if $\chi(\zeta) = (T-1)/[i(0)-1]$, then

$$w_r = \frac{2g[i(0)-1]x}{u_1} w(\zeta),$$

where $w(\zeta)$ is the solution of

$$\frac{d}{d\zeta} (\theta w') + \eta w' - 2\eta' w + \chi = 0, \quad (59)$$

subject to the boundary conditions $w(0) = w(\infty) = 0$. The required solution is that linear combination of the solutions of the equation with the boundary conditions (i) $w(0) = w'(0) = 0$ and (ii) $w(0) = 0, w'(0) = 1$ that gives $w(\infty) = 0$.

The computation of w has been carried out in the easiest case, when $\sigma = 1$ and $\theta = 1$, the plate being heat-insulated. The function η then satisfies equation (39), and

$$\eta(\zeta) = \frac{1}{\sqrt{2}}f_0\left(\frac{\zeta}{\sqrt{2}}\right), \quad \eta'(\zeta) = \frac{1}{2}f'_0\left(\frac{\zeta}{\sqrt{2}}\right),$$

where $f_0(\xi)$ satisfies

$$f_0''' + f_0 f_0'' = 0, \quad f_0(0) = f_0'(0) = 0, \quad f_0'(\infty) = 2.$$

The functions $f_0(\xi), f_0'(\xi)$ are tabulated, for example, by Howarth (1938). Also, $i + \frac{1}{2}bu^2 = 1 + \frac{1}{2}b$, so $\chi = 1 - \eta'^2$. The values of w are given in table 2. It is important to notice that the maximum drift velocity across any section x is proportional to the distance from the leading edge.

TABLE 2. THE VELOCITY COMPONENTS IN THE BOUNDARY LAYER OF A HEAT-INSULATED VERTICAL FLAT PLATE, WHEN $\mu_r \propto T_r, \sigma = 1$

$\xi = \zeta/\sqrt{2}$	$u = \frac{1}{2}f'_0(\xi)$	w
0	0	0
0.1	0.061	0.099
0.2	0.133	0.178
0.3	0.199	0.238
0.4	0.265	0.280
0.5	0.330	0.307
0.6	0.394	0.319
0.7	0.456	0.320
0.8	0.517	0.310
0.9	0.575	0.293
1.0	0.630	0.270
1.1	0.681	0.244
1.2	0.729	0.216
1.3	0.772	0.187
1.4	0.812	0.159
1.5	0.846	0.133
1.6	0.876	0.109
1.7	0.902	0.087
1.8	0.923	0.069
1.9	0.941	0.054
2.0	0.956	0.041
2.5	0.992	0.008
3.0	0.999	0.000

In the case of a vertical blunt-nosed cylinder the values of η and i have been computed for $u_1 = \beta_1 x$ and are tabulated in table 1. Using these values together with the assumption $i = T$ and the same notation for $\chi(\zeta)$ as in the previous case, the solution of (58) is

$$w_r = \frac{2g[i(0) - 1]}{\beta_1} w(\zeta),$$

where $w(\zeta)$ satisfies

$$\frac{d}{d\zeta}(\theta w') + \eta w' + \frac{1}{2}\chi = 0 \quad (60)$$

and $w(0) = w(\infty) = 0$. The solution is tabulated in table 1. Here the maximum drift velocity across a section is independent of the distance from the foremost generator.

The values of w , having been determined, it is possible to check *a posteriori* the assumption that the drift has only a small influence on the horizontal primary flow. The interaction of the drift and the primary flow occurs only in the energy equation (13), which may be written, after transformation,

$$\frac{i_1 \phi_1}{\sigma} \frac{\partial}{\partial \psi} \left(\phi \frac{\partial i}{\partial \psi} \right) - i_1 \frac{\partial i}{\partial x} + \phi_1 u_1^2 \phi \left(\frac{\partial u}{\partial \psi} \right)^2 + \phi_1 w_1^2 \phi \left(\frac{\partial w}{\partial \psi} \right)^2 - \frac{g w_1}{u_1} \frac{i w}{u} = 0, \quad (61)$$

where $w_1 = \frac{w_r}{w} = 2g[i(0) - 1]x/u_1$ or $2g[i(0) - 1]/\beta_1$.

The terms containing w , and previously neglected, are the last two terms on the left.

For a flat plate, with the notation of § 3, equation (61) may be transformed into

$$\frac{1}{\sigma} \frac{d}{d\zeta} (\theta i') + \eta i' + b \theta \eta'^2 + \frac{w_1^2}{i_1} w'^2 - \frac{2g w_1 x}{u_1 i_1} i w = 0,$$

the last two terms being the terms neglected in § 3.

Since
$$\frac{w_1^2}{i_1} = (\gamma - 1) \left\{ \frac{2gx}{M_1 a_1^2} [i(0) - 1] \right\}^2$$

and
$$\frac{g w_1 x}{u_1 i_1} = 2(\gamma - 1) [i(0) - 1] \left(\frac{gx}{M_1 a_1^2} \right)^2,$$

and these expressions are both small compared with unity except for large values of x , comparable with $M_1 a_1^2/g$, it follows that a good approximation will be obtained (except for large x) if the horizontal primary flow for a vertical plate is calculated with the vertical drift neglected. The same result is true for a sharp-nosed cylinder.

Near the front of a blunt-nosed cylinder, with the notation of § 4 and with $u_1 = \beta_1 x$, the energy equation becomes

$$\eta' \left[\frac{1}{\sigma} \frac{d}{d\zeta} (\theta i') + \eta i' \right] + \frac{w_1^2}{i_{10}} w'^2 - \frac{g w_1}{\beta_1 i_{10}} \eta' i w = 0,$$

when, as in § 4, the dissipation due to the horizontal primary flow is neglected. The last two terms on the left were neglected in § 4. Since

$$\frac{w_1^2}{i_{10}} = (\gamma - 1) \left\{ \frac{2g[i(0) - 1]}{\beta_1 a_{10}} \right\}^2 \quad \text{and} \quad \frac{g w_1}{\beta_1 i_{10}} = 2(\gamma - 1) [i(0) - 1] \left(\frac{g}{\beta_1 a_{10}} \right)^2,$$

the neglect of these last two terms is justified if $g/(\beta_1 a_{10})$ is small. If we accept the estimate of β_1 in § 2.31, this condition becomes $2g/[3a_1^2(1 + \frac{1}{2}b)^{\frac{1}{2}}\kappa] \ll 1$, so, if the neglect of the dissipation due to the horizontal primary flow is justified, the results of § 4 will provide a good approximation near the front of a blunt-nosed vertical cylinder, unless its radius of curvature at the nose is very large.

It is probable that the influence of drift on the flow in the x -direction may be neglected also for a sloping blunt-nosed cylinder, unless its radius of curvature at the nose is very large. For such a cylinder the secondary flow in the x -direction will be negligible, when the cross-drift is, only if the stream is vertical.

7. A MORE ACCURATE SOLUTION FOR A HORIZONTAL BLUNT-NOSED CYLINDER IN A VERTICAL STREAM

In § 2 it was shown that for a horizontal blunt-nosed cylinder the solution in § 4 for the primary flow near the forward stagnation point is a good approximation to the total flow only if the stream is vertical and the cylinder is not too blunt. In this particular case it is possible fairly easily to solve the boundary-layer equations, viz. (14), (16), (18) and (19) of § 3, if dissipation is neglected, and c_p is constant, so that $i = T$. The exact equations become, after transformation,

$$u_1 \frac{du_1}{dx} (u^2 - T) + u_1^2 u \frac{\partial u}{\partial x} = gl(T-1) + \phi_1 u_1^2 u \frac{\partial}{\partial \psi} \left(\phi \frac{\partial u}{\partial \psi} \right), \quad (62)$$

$$\frac{di_1}{dx} (i - T) + i_1 \frac{\partial i}{\partial x} = \frac{\phi_1 i_1}{\sigma} \frac{\partial}{\partial \psi} \left(\phi \frac{\partial i}{\partial \psi} \right) + \phi_1 u_1^2 \phi \left(\frac{\partial u}{\partial \psi} \right)^2, \quad (63)$$

and

$$\frac{di_1}{dx} + gl + u_1 \frac{du_1}{dx} = 0.$$

The equations may be reduced to ordinary differential equations if $u_1^2 \propto lx$, $l \propto x^\lambda$, and θ_1 may be taken as constant. If $l \propto x$, $u_1 \propto x$, the conditions are those that occur near the foremost generator of a blunt-nosed cylinder in a vertical stream (§ 2). Then $l \doteq \kappa x$, $u_1 = \beta_1 x$, and, as in § 4, $i_1 = i_{10} + O(u_1^2)$ and $\theta_1 = \theta_{10} + O(x^2)$, the suffix 10 referring to the forward stagnation point. When $i = T$ and dissipation is neglected (as in § 4), equations (62) and (63) reduce to

$$\left. \begin{aligned} \frac{d}{d\zeta} (\theta \eta'') + \eta \eta'' - \eta'^2 + i + \frac{g\kappa}{\beta_1^2} (i-1) &= 0, \\ \frac{d}{d\zeta} (\theta i') + \sigma \eta i' &= 0. \end{aligned} \right\} \quad (64)$$

When $g\kappa/\beta_1^2 \ll 1$, and is neglected, these equations are the same as (41) and (42) for $m = \frac{1}{2}$, as they should be (see condition (21)). The equations (64) may be solved numerically (for any given value of $g\kappa/\beta_1^2$) by a process similar to that used in § 4 for equations (41) and (42).

8. FREE CONVECTION CURRENTS AT A FLAT PLATE

The exact equations (62) and (63) of § 7 govern the flow past a flat plate whose leading edge is horizontal. In this case l is constant, and the method of § 7 may be applied only if $u_1 \propto x^{\frac{1}{2}}$, which is not appropriate unless $u_1 = 0$, u_1 being the velocity of the main stream outside the boundary layer. We consider further the case of free convection, when the surrounding fluid is at rest. It will be assumed that c_p is constant, so that $i = T$, and the notation will be changed, u_1 now being used for the representative velocity $2\{gl[i(0)-1]x\}^{\frac{1}{2}}$, with $u_r = u_1 u$. The equations governing the free convective flow may then be written

$$u_1 u \frac{\partial}{\partial x} (u_1 u) = gl(T-1) + \phi_1 u_1^2 \frac{\partial}{\partial \psi} \left(\phi \frac{\partial u}{\partial \psi} \right), \quad (65)$$

$$i_1 \frac{\partial i}{\partial x} = \frac{\phi_1 i_1}{\sigma} \frac{\partial}{\partial \psi} \left(\phi \frac{\partial i}{\partial \psi} \right) + \phi_1 u_1^2 \phi \left(\frac{\partial u}{\partial \psi} \right)^2, \quad (66)$$

and

$$\frac{di_1}{dx} + gl = 0.$$

The flow to which these equations refer is produced by the difference in temperature between the plate and the surrounding fluid. The equations apply only to a flow in a boundary layer, restricted to a neighbourhood of a fixed boundary. Consequently for a vertical plate ($l = 1$) they govern the free convection current on each side, but for a sloping plate they apply only to the upper or to the lower side according as the plate is cold or hot relative to the fluid.

The enthalpy i_1 of the surrounding fluid varies with altitude, and

$$i_1 = i_{10} - glx,$$

where the suffix 10 refers to the fluid at the level of the bottom edge of the plate. The fractional error in neglecting this variation and taking $i_1 = i_{10}$ is glx/i_{10} , which for air at atmospheric temperature and pressure is $lx/980$ % at a distance x ft. up the plate. For distances less than, say, 10 ft. in air, it will be a good approximation to take $i_1 = i_{10}$, and therefore $\theta_1 = \theta_{10}$. Near the bottom of the plate the velocity of the flow will also be small, and the heat of dissipation may be neglected in (66). In these circumstances (65) and (66) are reducible to

$$\frac{d}{d\xi}(\theta\eta'') + 3\eta\eta'' - 2\eta'^2 + \chi = 0, \quad (67)$$

$$\frac{d}{d\xi}(\theta\chi') + 3\sigma\eta\chi' = 0, \quad (68)$$

where $u_1 = 2\sqrt{\{gl[i(0) - 1]x\}}$, $\eta = \psi\{64\theta_{10}^2 gl[i(0) - 1]x^3\}^{-1/4}$,

$$\chi = \frac{i_r - i_1}{i_0 - i_1} = \frac{T - 1}{i(0) - 1} \quad \text{and} \quad \xi = \int_0^\eta \frac{d\tau}{u(\tau)},$$

so that $u = \eta'$, as before. The boundary conditions are

$$\eta(0) = \eta'(0) = 0, \quad \eta'(\infty) = 0, \quad \chi(0) = 1 \quad \text{and} \quad \chi(\infty) = 0.$$

If $\theta = 1$, (67) and (68) reduce to equations solved by Pohlhausen (see Schmidt 1934). Starting with Pohlhausen's solution, when the accurate expression for θ is used, u and χ may be found by successive approximations from the equations

$$u(\xi) = \frac{K(\infty)}{J(\infty)} J(\xi) - K(\xi), \quad (69)$$

$$\chi(\xi) = 1 - L(\xi)/L(\infty), \quad (70)$$

where $J(\xi) = \int_0^\xi \frac{E^3}{\theta} d\xi$, $K(\xi) = \int_0^\xi \frac{E^3}{\theta} \left(\int_0^\xi \frac{\chi - 2\eta'^2}{E^3} d\xi \right) d\xi$,

and $L(\xi) = \int_0^\xi \frac{E^{3\sigma}}{\theta} d\xi$.

There is a difficulty in computing K because $\chi - 2\eta'^2$ and E both tend to zero as $\xi \rightarrow \infty$. A similar difficulty was encountered previously in § 4 over the integral I . Although no calculations have been made in this case of free convection, it seems

likely that the procedure used for the calculations described in § 4 would be suitable here too.

If Nu is the Nusselt number of the flow and Gr the Grasshoff number,

$$Nu = \frac{2\sqrt{2}}{3} \theta(0) \chi'(0) (Gr)^{\frac{1}{2}},$$

where

$$Gr = \frac{gh^3[i(0) - 1]}{\nu_1^2},$$

and h is the distance measured up the plate parallel to the x -axis from the leading edge.

There are cases in which the refinement produced by using the accurate law of viscosity instead of $\theta = 1$ is unimportant. For example, in the experiments of Schmidt (1934) with a vertical plate at 65°C in air at 15°C , the values of θ vary between 0.964 at the plate and 1 at the outside edge of the boundary layer. The error incurred by using Pohlhausen's values for $\eta(\zeta)$ and $\chi(\zeta)$ can hardly be more than 2% in this case.

9. FLOW IN PLANE JETS AND WAKES

So far as the flow in a plane jet remains laminar, it may be assumed that the flow is governed by the laminar boundary-layer equations, and that, except near the orifice, the pressure is uniform. If the plane of symmetry is $y = 0$, the equations for a jet in which the secondary flow is negligible are

$$\frac{\partial}{\partial x}(\rho_r u_r) + \frac{\partial}{\partial y}(\rho_r v_r) = 0, \quad (71)$$

$$\rho_r \left(u_r \frac{\partial u_r}{\partial x} + v_r \frac{\partial u_r}{\partial y} \right) = \frac{\partial}{\partial y} \left(\mu_r \frac{\partial u_r}{\partial y} \right), \quad (72)$$

$$\rho_r \left(u_r \frac{\partial i_r}{\partial x} + v_r \frac{\partial i_r}{\partial y} \right) = \frac{1}{\sigma} \frac{\partial}{\partial y} \left(\mu_r \frac{\partial i_r}{\partial y} \right) + \mu_r \left(\frac{\partial u_r}{\partial y} \right)^2. \quad (73)$$

At $y = 0$, by symmetry $v_r = 0$, $\partial u_r / \partial y = 0$ and $\partial i_r / \partial y = 0$. At the boundaries of the jet $u_r = 0$ and $i_r = i_1 = \text{const}$. The value of u_r on the axis, $y = 0$, will be denoted by u_0 , and we write $u_r = u_0 u$, $i_r = i_1 i$.

By integrating (71), (72) and (73) with respect to y it follows that

$$\frac{d}{dx} \int_0^\infty \rho_r u_r^2 dy = 0, \quad (74)$$

and

$$\frac{d}{dx} \int_0^\infty \rho_r u_r (i_r + \frac{1}{2} u_r^2 - i_1) dy = 0. \quad (75)$$

In terms of the independent variables x and ψ , the equations (72) and (73) are

$$\left(\frac{d}{dx} \ln u_0 \right) u + \frac{\partial u}{\partial x} = \phi_1 \frac{\partial}{\partial \psi} \left(\phi \frac{\partial u}{\partial \psi} \right), \quad (76)$$

$$\frac{\partial i}{\partial x} = \frac{\phi_1}{\sigma} \frac{\partial}{\partial \psi} \left(\phi \frac{\partial i}{\partial \psi} \right) + \phi_1 b \phi \left(\frac{\partial u}{\partial \psi} \right)^2, \quad (77)$$

where $\phi_1 = \theta_1 u_0$ and $b = u_0^2/i_1$. Let us investigate the conditions under which (76) and (77) may be reduced to ordinary differential equations, i.e. the conditions that u , i and ϕ should be functions of a single variable $\eta = \psi/f(x)$. Then

$$\int_0^\infty \rho_r u_r^2 dy = \int_0^\infty u_r d\psi = \frac{u_0(x)}{f(x)} \int_0^\infty u(\eta) d\eta,$$

$$\int_0^\infty \rho_r u_r (i_r + \frac{1}{2} u_r^2 - i_1) dy = \frac{i_1}{f(x)} \int_0^\infty [i(\eta) - 1] d\eta + \frac{u_0^2(x)}{2f(x)} \int_0^\infty u^2(\eta) d\eta.$$

The first expression cannot be independent of x (as required by (74)) unless u_0/f is constant, and then the second expression cannot be independent of x (see (75)) unless it is identically zero. The latter requirement is met if we assume that

$$i_r + \frac{1}{2} u_r^2 - i_1 = 0, \quad (78)$$

and this is equivalent to assuming that $\sigma = 1$. Equation (77) is now replaced by (78), but (76) is still not reducible to an ordinary differential equation unless $\theta = 1$, since by (78) i is not a function of η only. To reduce (76) and (77) to ordinary differential equations the limitations that $\sigma = 1$ and $\theta = 1$ have therefore been introduced and these are just the limitations used by Howarth (1948) in his discussion of the plane jet. With these limitations $u_0 = \delta/x^{\frac{1}{2}}$, where δ is a constant, and with

$$\eta = \frac{\psi}{(6\theta_1 \delta)^{\frac{1}{2}} x^{\frac{1}{2}}}$$

equation (76) becomes
$$\frac{d}{d\eta} \left(u \frac{du}{d\eta} \right) + 2 \left(\eta \frac{du}{d\eta} + u \right) = 0, \quad (79)$$

the boundary condition being $(du/d\eta)_{\eta=0} = 0$ and $u(0) = 1$. The solution is

$$u = 1 - \eta^2.$$

The relation between δ and the total flux of momentum M across any section of the jet is

$$M = \int_{-\infty}^{\infty} \rho_r u_r^2 dy = (6\theta_1 \delta^3)^{\frac{1}{2}} \int_{-1}^1 u d\eta = 4(\frac{2}{3}\theta_1 \delta^3)^{\frac{1}{2}}.$$

The velocity and enthalpy in the jet are given by

$$u_r = \left(\frac{3M^2}{32\theta_1 x} \right)^{\frac{1}{2}} (1 - \eta^2), \quad i_r = i_1 - \frac{1}{2} u_r^2. \quad (81)$$

The y co-ordinate is given by

$$y = \int_0^\psi \frac{d\psi}{\rho_r u_r} = \left(\frac{6\nu_1}{\delta} \right)^{\frac{1}{2}} x^{\frac{1}{2}} \int_0^\eta \frac{T}{u} d\eta.$$

When c_p is constant, so that $T = i$,

$$y = \left(\frac{6\nu_1}{\delta} \right)^{\frac{1}{2}} x^{\frac{1}{2}} \int_0^\eta \left[\frac{1}{1 - \eta^2} - \frac{\delta^2}{2i_1 x^{\frac{1}{2}}} (1 - \eta^2) \right] d\eta$$

$$= \left(\frac{48\mu_1 \nu_1}{M} x^2 \right)^{\frac{1}{2}} \tanh^{-1} \eta - \frac{3M}{8\rho_1 i_1} \eta (1 - \frac{1}{2} \eta^2).$$

Laminar flow in a plane wake is also assumed to be governed by the equations (71), (72) and (73) of a laminar boundary layer with uniform pressure. Outside the wake, which is symmetrical about $y = 0$, the velocity is uniform and is denoted by u_1 ; inside the wake it is put equal to $u_r = u_1(1 - \Delta u_r)$, where Δu_r is small. Similarly, inside the wake, $i_r = i_1(1 - \Delta i_r)$, where Δi_r is small. It is assumed that secondary flow may be neglected.

The equations

$$\frac{d}{dx} \int_0^\infty \rho_r u_r (u_r - u_1) dy = 0, \quad (82)$$

$$\frac{d}{dx} \int_0^\infty \rho_r u_r (i_r + \frac{1}{2} u_r^2 - i_1 - \frac{1}{2} u_1^2) dy = 0 \quad (83)$$

must be satisfied. If u_r , ρ_r , i_r , etc., are functions of a single independent variable $\eta = \psi f(x)$, then

$$\int_0^\infty \rho_r u_r (u_r - u_1) dy = \int_0^\infty (u_r - u_1) d\psi = -\frac{u_1}{f(x)} \int_0^\infty \Delta u_r d\eta,$$

and

$$\begin{aligned} \int_0^\infty \rho_r u_r (i_r + \frac{1}{2} u_r^2 - i_1 - \frac{1}{2} u_1^2) dy &= \int_0^\infty (i_r + \frac{1}{2} u_r^2 - i_1 - \frac{1}{2} u_1^2) d\psi \\ &= -\frac{1}{f(x)} \int_0^\infty (i_1 \Delta i_r + u_1^2 \Delta u_r) d\eta, \end{aligned}$$

when $(\Delta u_r)^2$ is neglected. These two expressions are independent of x , and (82) and (83) are satisfied, if $\Delta u_r = f(x) \Delta u$, $\Delta i_r = f(x) \Delta i$, and Δu , Δi are functions of η . Equations (71), (72) and (73) may then be transformed into

$$\frac{\partial}{\partial x} [f(x) \Delta u] = \phi_1 f(x) \frac{\partial^2}{\partial \psi^2} \Delta u, \quad \frac{\partial}{\partial x} [f(x) \Delta i] = \frac{\phi_1}{\sigma} f(x) \frac{\partial^2}{\partial \psi^2} \Delta i,$$

terms of the second order of small quantities being neglected, and since Δu , Δi are functions of η we may take $f(x) = \frac{1}{2}(\phi_1 x)^{-\frac{1}{2}}$. The equations then become

$$\frac{d^2}{d\eta^2} \Delta u + 2 \left(\eta \frac{d}{d\eta} \Delta u + \Delta u \right) = 0, \quad \frac{d^2}{d\eta^2} \Delta i + 2\sigma \left(\eta \frac{d}{d\eta} \Delta i + \Delta i \right) = 0. \quad (84)$$

By symmetry, $\frac{d}{d\eta} \Delta u = \frac{d}{d\eta} \Delta i = 0$ when $\eta = 0$, and the required solutions of (84) are

$$\Delta u = A e^{-\eta^2}, \quad \Delta i = B e^{-\sigma \eta^2},$$

where A and B are constants.

Assuming that the pressure of the gas is the same upstream and downstream, the drag D per unit width of the cylindrical obstacle producing the wake is given by

$$D = \int_{-\infty}^\infty \rho_r u_r (u_r - u_1) dy = 2u_1 \int_0^\infty \Delta u d\eta = \sqrt{\pi} A u_1$$

If $-2Q_0$ is the rate of transfer of heat from the gas to the cylinders,

$$-2Q_0 = 2i_1 \int_0^\infty \Delta i d\eta,$$

i.e.

$$Q_0 = -B i_1 \int_0^\infty e^{-\sigma \eta^2} d\eta = -\frac{1}{2} \left(\frac{\pi}{\sigma} \right)^{\frac{1}{2}} B i_1.$$

The velocity and enthalpy in the wake are, therefore,

$$u_r = u_1 - \frac{1}{2}D(\pi\phi_1 x)^{-\frac{1}{2}}e^{-\eta^2}, \quad i_r = i_1 + Q_0(\pi\phi_1 x/\sigma)^{-\frac{1}{2}}e^{-\sigma\eta^2}. \quad (85)$$

The y co-ordinate is given by

$$y = \frac{2(\phi_1 x)^{\frac{1}{2}}}{\rho_1 u_1} \int_0^\eta \frac{T_r/T_1}{1 - \Delta u_r} d\eta.$$

Since the variation in T_r is small, c_p may be taken as constant and

$$T_r/T_1 = i_r/i_1 = 1 - \Delta i_r.$$

Consequently,

$$\begin{aligned} y &= 2\left(\frac{\nu_1 x}{u_1}\right)^{\frac{1}{2}} \int_0^\eta (1 + \Delta u_r - \Delta i_r) d\eta \\ &= 2\left(\frac{\nu_1 x}{u_1}\right)^{\frac{1}{2}} \eta + \frac{D}{2\rho_1 u_1^2} \operatorname{erf} \eta + \frac{Q_0}{\rho_1 u_1 i_1} \operatorname{erf} \sqrt{\sigma} \eta. \end{aligned} \quad (86)$$

When the obstacle is completely surrounded by its own laminar boundary layer and the drag on it is entirely due to skin friction, equations (85) and (86) can be written conveniently in terms of non-dimensional parameters. Consider, for example, the wake of a flat plate of length L in a uniform stream, for which (cf. § 3)

$$D = 2 \int_0^L \mu_0 \left(\frac{\partial u_r}{\partial y} \right)_{y=0} dx, \quad Q_0 = \int_0^L q_0 dx,$$

when the suffix 0 refers to conditions at the plate. In § 3, $\eta = \psi/(2\phi_1 x)^{\frac{1}{2}}$, and with this definition of η ,

$$\frac{D}{2\rho_1 u_1^2} = \left(\frac{2}{R}\right)^{\frac{1}{2}} L \left(\phi \frac{du}{d\eta} \right)_{\eta=0},$$

and

$$\frac{Q_0}{\rho_1 u_1 i_1} = -\frac{1}{\sigma} \left(\frac{2}{R}\right)^{\frac{1}{2}} L \left(\phi \frac{di}{d\eta} \right)_{\eta=0},$$

where R is the Reynolds number $u_1 L/\nu_1$. The drag coefficient $c_f = \frac{D}{\rho_1 u_1^2 L}$, and the

Nusselt number $Nu = \frac{\sigma Q_0}{\mu_1(i_0 - i_1)}$, so $\frac{Q_0}{\rho_1 u_1 i_1 L} = \frac{1}{\sigma R} [i(0) - 1] Nu$.

Hence, returning to the definition of η used in this section, namely, $\eta = \frac{1}{2}\psi/(\phi_1 x)^{\frac{1}{2}}$, we have

$$\frac{u_r}{u_1} = 1 - \frac{1}{2} \left(\frac{RL}{\pi x} \right)^{\frac{1}{2}} c_f e^{-\eta^2},$$

$$\frac{i_r}{i_1} = 1 + \left(\frac{L}{\pi \sigma R x} \right)^{\frac{1}{2}} [i(0) - 1] Nu e^{-\sigma \eta^2},$$

$$\frac{y}{L} = 2 \left(\frac{x}{RL} \right)^{\frac{1}{2}} \eta + \frac{1}{2} c_f \operatorname{erf} \eta + \frac{1}{\sigma R} [i(0) - 1] Nu \operatorname{erf} \sqrt{\sigma} \eta.$$

I wish to express my sincere thanks to Professor S. Goldstein for various suggestions in connexion with this paper.

REFERENCES

- Cope, W. F. & Hartree, D. R. 1948 *Phil. Trans. A*, **241**, 1.
Crocco, L. 1946 *Monogr. Sci. Aero.* no. 3.
Ferri, A. 1942 *Atti Guidonia*, nos. 67, 68, 69.
Goldstein, S. (editor) 1938 *Modern developments in fluid dynamics*. Oxford: Clarendon Press.
Hartree, D. R. 1949 *Rep. Memor. Aero. Res. Comm., Lond.*, no. 2426.
Howarth, L. 1934 *Rep. Memor. Aero. Res. Comm., Lond.*, no. 1632.
Howarth, L. 1938 *Proc. Roy. Soc. A*, **164**, 547.
Howarth, L. 1948 *Proc. Roy. Soc. A*, **194**, 16.
Kármán, Th. von & Tsien, H. S. 1938 *J. Aero. Sci.* **5**, 227.
Pohlhausen, E. 1921 *Z. angew. Math. Mech.* **1**, 115.
Schmidt, E. 1934 *Proc. 4th Int. Congr. Appl. Mech.* p. 98.

INDEX TO VOLUME 199 (A)

- Absorption and dispersion of ultrasonic waves in acetic acid (Lamb & Pinkerton), 114.
Ambrose, E. J., Elliott, A. & Temple, R. B. The use of polarized infra-red radiation in the study of doubly oriented long-chain polymers, 183.
Analytic behaviour of Heisenberg's S matrix (Eden), 256.
Angular distribution of γ -radiation from light nuclei. I. Experimental (Devons & Hine), 56.
Angular distribution of γ -radiation from light nuclei. II. Theoretical (Devons & Hine), 73.
Bakerian Lecture. A region of biosynthesis (Raistrick), 141.
Batchelor, G. K. & Townsend, A. A. The nature of turbulent motion at large wave-numbers, 238.
Binnie, A. M. & Davidson, J. F. The flow under gravity of a swirling liquid through an orifice-plate, 443.
Bowden, K. F. & Proudman, J. Observations on the turbulent fluctuations of a tidal current, 311.
Bubble rafts producing optical diffraction patterns (Dyson), 130.
Bullard, E. C. Electromagnetic induction in a rotating sphere, 413.
Burden, F. A. & Burgoyne, J. H. The ignition and flame reactions of ethylene oxide, 328.
Burgoyne, J. H. *See* Burden & Burgoyne.
Cochrane, W. & Hester, A. G. The transformation of the nitrogen isotope of mass 15 into carbon and helium by bombardment with protons, 458.
Conjugated hydrocarbons, structure (Moffitt), 487.
Cottrell, A. H. & Jaswon, M. A. Distribution of solute atoms round a slow dislocation, 104.
Davidson, J. F. *See* Binnie & Davidson.
Davies, T. V. Unsteady compressible flow in two dimensions, 468.
Devons, S. & Hine, M. G. N. The angular distribution of γ -radiation from light nuclei. I. Experimental, 56.
Devons, S. & Hine, M. G. N. The angular distribution of γ -radiation from light nuclei. II. Theoretical, 73.
Distribution of solute atoms round a slow dislocation (Cottrell & Jaswon), 104.
Domb, C. Order-disorder statistics. II. A two-dimensional model, 199.
Dyson, J. Optical diffraction patterns produced by bubble rafts, 130.
Eden, R. J. The analytic behaviour of Heisenberg's S matrix, 256.
Elastic theory of dislocations (Mann), 376.
Electromagnetic induction in a rotating sphere (Bullard), 413.
Electron pairing theory of the structure of conjugated hydrocarbons (Moffitt), 487.
Elliott, A. *See* Ambrose, Elliott & Temple.
Ethylene oxide, ignition and flame reactions (Burden & Burgoyne), 328.
Flow under gravity of a swirling liquid through an orifice-plate (Binnie & Davidson), 443.
Gaydon, A. G. & Wolfhard, H. G. Spectroscopic studies of low-pressure flames. II. Effective translational and rotational temperatures from CH bands, 89.
Grassie, N. & Melville, H. W. The thermal degradation of polyvinyl compounds. I. A new type of molecular still, 1.

- Grassie, N. & Melville, H. W. The thermal degradation of polyvinyl compounds. II. The degradation of benzoyl peroxide catalyzed polymethyl methacrylates, 14.
- Grassie, N. & Melville, H. W. The thermal degradation of polyvinyl compounds. III. The effect of inhibitors and end groups on the degradation of polymethyl methacrylate, 24.
- Grassie, N. & Melville, H. W. The thermal degradation of polyvinyl compounds. IV. The thermal degradation of the methyl methacrylate copolymers with glycol dimethacrylate and acrylonitrile, 39.
- Hatton, J. & Rollin, B. V. Nuclear magnetic resonance at low temperatures, 222.
- Heisenberg's *S* matrix, analytic behaviour of (Eden), 256.
- Helium, transition of liquid (Temperley), 361.
- Hester, A. G. *See* Cochrane & Hester.
- Hillson, P. J. & Rideal, E. K. On overpotential and the photovoltaic process at polarized electrodes, 295.
- Hine, M. G. N. *See* Devons & Hine.
- Ignition and flame reactions of ethylene oxide (Burden & Burgoyne), 328.
- Illingworth, C. R. Steady flow in the laminar boundary layer of a gas, 533.
- Jaswon, M. A. *See* Cottrell & Jaswon.
- Kilmister, C. W. The use of quaternions in wave-tensor calculus, 517.
- Kinetics of the polymerization of sarcosine carbonic anhydride (Waley & Watson), 499.
- Lamb, J. & Pinkerton, J. M. M. The absorption and dispersion of ultrasonic waves in acetic acid, 114.
- Long-chain polymers, study of doubly oriented (Ambrose, Elliott & Temple), 183.
- Mann, E. H. An elastic theory of dislocations, 376.
- Marston, H. R. The organization and work of the Division of Biochemistry and General Nutrition of C.S.I.R., 273.
- Melville, H. W. *See* Grassie & Melville.
- Moffitt, W. The electron pairing theory of the structure of conjugated hydrocarbons, 487.
- Nature of turbulent motion at large wave-numbers (Batchelor & Townsend), 238.
- Nuclear magnetic resonance at low temperatures (Hatton & Rollin), 222.
- Observations on the turbulent fluctuations of a tidal current (Bowden & Proudman), 311.
- Optical diffraction patterns produced by bubble rafts (Dyson), 130.
- Order-disorder statistics. II. A two-dimensional model (Domb), 199.
- Organization and work of the Division of Biochemistry and General Nutrition of C.S.I.R. (Marston), 273.
- Overpotential and the photovoltaic process at polarized electrodes (Hillson & Rideal), 295.
- Pinkerton, J. M. M. *See* Lamb & Pinkerton.
- Polyvinyl compounds, thermal degradation (Grassie & Melville), 1, 14, 24, 39.
- Production of mesons by electrons (Sneddon & Touschek), 352.
- Proudman, J. *See* Bowden & Proudman.
- Pyrolysis of methane, ethane and *n*-butane on a platinum filament (Robertson), 394.

- Raistrick, H. A region of biosynthesis (Bakerian Lecture), 141.
Region of biosynthesis (Bakerian Lecture) (Raistrick), 141.
Rideal, E. K. *See* Hillson & Rideal.
Robertson, A. J. B. The pyrolysis of methane, ethane and *n*-butane on a platinum filament, 394.
Rollin, B. V. *See* Hatton & Rollin.
Rotating sphere, electromagnetic induction (Bullard), 413.
- Sarcosine carbonic anhydride, polymerization (Waley & Watson), 499.
Simpson, D. M. & Sutherland, G. B. B. M. Vibration spectra of hydrocarbon molecules. II. Skeletal frequencies in certain branched paraffins, 169.
Sneddon, I. N. & Touschek, B. F. The production of mesons by electrons, 352.
Spectroscopic studies of low-pressure flames. II. Effective translational and rotational temperatures from CH bands (Gaydon & Wolfhard), 89.
Statistical mechanics and the partition of numbers. I. The transition of liquid helium (Temperley), 361.
Steady flow in the laminar boundary layer of a gas (Illingworth), 533.
Sutherland, G. B. B. M. *See* Simpson & Sutherland.
- Temperley, H. N. V. Statistical mechanics and the partition of numbers. I. The transition of liquid helium, 361.
Temple, R. B. *See* Ambrose, Elliott & Temple.
Thermal degradation of polyvinyl compounds. I. A new type of molecular still (Grassie & Melville), 1.
Thermal degradation of polyvinyl compounds. II. The degradation of benzoyl peroxide catalyzed polymethyl methacrylates (Grassie & Melville), 14.
Thermal degradation of polyvinyl compounds. III. The effect of inhibitors and end groups on the degradation of polymethyl methacrylate (Grassie & Melville), 24.
Thermal degradation of polyvinyl compounds. IV. The thermal degradation of the methyl methacrylate copolymers with glycol dimethacrylate and acrylonitrile (Grassie & Melville), 39.
Tidal current, turbulent fluctuations of a (Bowden & Proudman), 311.
Touschek, B. F. *See* Sneddon & Touschek.
Townsend, A. A. *See* Batchelor & Townsend.
Transformation of the nitrogen isotope of mass 15 into carbon and helium by bombardment with protons (Cochrane & Hester), 458.
Turbulent motion at large wave-numbers (Batchelor & Townsend), 238.
- Ultrasonic waves in acetic acid (Lamb & Pinkerton), 114.
Unsteady compressible flow in two dimensions (Davies), 468.
Use of polarized infra-red radiation in the study of doubly oriented long-chain polymers (Ambrose, Elliott & Temple), 183.
Use of quaternions in wave-tensor calculus (Kilmister), 517.
- Vibration spectra of hydrocarbon molecules. II. Skeletal frequencies in certain branched paraffins (Simpson & Sutherland), 169.
- Waley, S. G. & Watson, J. The kinetics of the polymerization of sarcosine carbonic anhydride, 499.
Watson, J. *See* Waley & Watson.
Wolfhard, H. G. *See* Gaydon & Wolfhard.

I.A.R.I. 75

INDIAN AGRICULTURAL RESEARCH
INSTITUTE LIBRARY, NEW DELHI.

[illegible]

GIPNLK—H-40 I.A.R.I.—29-4-55—15,000